DOCUMENT RESUME

ED 144 980 TM 006 530

AUTHOR Rest, James; And Others
TITLE Development in Judging Moral Issues--A Summary of Research Using the Defining Issues Test. Minnesota Moral Research Projects, Technical Report No. 3.
INSTITUTION Minnesota Univ., Minneapolis.
SPONS AGENCY National Inst. of Mental Health (DHEW), Rockville, Md.
PUB DATE 77
GRANT NIMH-R01-MH27861
NOTE 96p.; Tables will be marginally legible due to quality of print

EDRS PRICE MF-$0.83 HC-$4.67 Plus Postage.
DESCRIPTORS Adults; *Cognitive Development; College Students; *Correlation; Cross Sectional Studies; Delinquents; Higher Education; High School Students; Individual Characteristics; Individual Differences; Intervention; Junior High School Students; Longitudinal Studies; *Moral Development; Moral Values; Personality; Political Attitudes; Secondary Education; *Standardized Tests; Test Reliability; *Test Results; Test Validity
IDENTIFIERS *Defining Issues Test; *Moral Judgment

ABSTRACT

Six papers, originally presented in different form at the 1977 Convention of the Society for Research in Child Development, highlight approximately 100 studies that have used the Defining Issues Test (DIT) of moral judgment. The DIT's standardized format and objective scoring makes comparisons possible among studies, and this body of research comprises the most substantial data base yet accumulated on a single measure of moral judgment. The purpose of this report is to present outstanding examples of several types of DIT studies, to summarize the findings of similar studies, and to discuss the theoretical implications regarding the nature of moral judgment and cognitive-developmental theory. The papers include: Charting the course of development through cross-sectional and longitudinal studies (James R. Rest); Cognitive and political attitude correlates of the DIT (James L. Carroll); Moral judgment intervention studies using the DIT (Jeanette A. Lawrence); The DIT related to behavior in an experimental setting: promise keeping in the prisoner's dilemma game (M. Kathryn Jacobs); Social cognition related to behavior in a naturalistic setting: a comparison of delinquents, predelinquents and nondelinquents (Edgar B. McColgan); and Indexing moral development (Mark L. Davison and Stephen Robbins). (Author/MV)

Documents acquired by ERIC include many informal unpublished materials not available from other sources. ERIC makes every effort to obtain the best copy available. Nevertheless, items of marginal reproducibility are often encountered and this affects the quality of the microfiche and hardcopy reproductions ERIC makes available via the ERIC Document Reproduction Service (EDRS). EDRS is not responsible for the quality of the original document. Reproductions supplied by EDRS are the best that can be made from the original.

ED144980

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY James Rest

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

U S DEPARTMENT OF HEALTH EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

DEVELOPMENT IN JUDGING MORAL ISSUES--

A SUMMARY OF RESEARCH USING THE DEFINING ISSUES TEST

JAMES REST, UNIVERISTY OF MINNESOTA
JAMES CARROLL, ARIZONA STATE UNIVERSITY
JEANETTE LAWRENCE, UNIVERSITY OF MINNESOTA
KATHRYNE JACOBS, MEDICAL COLLEGE OF OHIO
EDGAR MCCOLGAN, UNIVERSITY OF TEXAS
MARK DAVISON, UNIVERSITY OF MINNESOTA
STEPHEN ROBBINS, UNIVERSITY OF MINNESOTA

From the 1977 Convention of the Society for Research in Child Development

MINNESOTA MORAL RESEARCH PROJECTS, TECHNICAL REPORT #3

This research was supported by a grant from the National Institutes of Mental Health, #R01 MH27861.

TM006 530

Development in Judging Moral Issues--

A Summary of Research Using the Defining Issues Test

James Rest, University of Minnesota
James Carroll, Arizona State University
Jeanette Lawrence, University of Minnesota
Kathryne Jacobs, Medical College of Ohio
Edgar McColgan, University of Texas
Mark Davison, University of Minnesota
Stephen Robbins, University of Minnesota

From the 1977 Convention of the Society for Research in Child Development

Minnesota Moral Research Projects, Technical Report #3

This research was supported by a grant from the National Institutes of Mental Health, #R01 MH27861.

CONTENTS

	Page
Introduction	1
Charting the Course of Development through Cross-sectional and Longitudinal Studies James R. Rest	7
Cognitive and Political Attitude Correlates of the Defining Issues Test James L. Carroll	18
Moral Judgment Intervention Studies using the Defining Issues Test Jeanette A. Lawrence	34
The DIT Related to Behavior in an Experimental Setting: Promise Keeping in the Prisoner's Dilemma Game M. Kathryne Jacobs	48
Social Cognition Related to Behavior in a Naturalistic Setting: A Comparison of Delinquents, Predelinquents and Nondelinquents Edgar B. McColgan	60
Indexing Moral Development Mark L. Davison and Stephen Robbins	70
Addendum: Personality Correlates of the DIT	78
References	82

Development in Judging Moral Issues--

A Summary of Research Using the Defining Issues Test

Introduction

The papers of this report give the highlights of about 100 studies that have used the Defining Issues Test of moral judgment. The papers were originally presented in a somewhat different form at the 1977 Convention of the Society for Research in Child Development. Currently about 100 studies using the Defining Issues Test (DIT) have been completed and several hundred more are in progress in most states of the U.S. and a number of foreign countries. The DIT's standardized format and objective scoring makes comparisons possible among studies, and this body of research comprises the most substantial data base yet accumulated on a single measure of moral judgment. The purpose of this report is to present outstanding examples of several types of DIT studies, to summarize the findings of similar studies, and to discuss the theoretical and practical implications regarding the nature of moral judgment and cognitive-developmental theory. The convergences and replications of the many studies have been remarkable. To my knowledge, this body of research represents the most thorough investigation of the nomological network of relationships yet conducted of the moral judgment construct, and no other measure of moral judgment has demonstrated such consistently high reliability and validity, involving so many different investigators with so many different populations.

The first paper summarizes cross-sectional studies (on about 6000 subjects) and 3 longitudinal studies, presenting evidence for developmental trends in the way people define moral issues (on the DIT). The second paper

by James Carroll reviews correlational studies relating to three questions: What is the evidence that the DIT is related to cognitive development? What is the relationship of the DIT to Kohlberg's test? What is the evidence that the DIT is related to values and social-political decisions? The third paper by Jeanette Lawrence reviews 14 educational intervention studies using the DIT as a pre-post measure of change, and points out desirable features of an intervention study. The fourth paper by Kathryne Jacobs relates the DIT to behavior in an experimental setting, and traces the linkages between the DIT's questionnaire responses and how subjects actually behave. The fifth paper by Edgar McColgan relates the DIT to behavior in a naturalistic setting, namely delinquent behavior, and compares the discriminative power of the DIT with Kohlberg's test and other social cognition measures. The sixth paper by Mark Davison presents a new procedure for scaling DIT responses and a new way of representing subjects' scores which increases the power of the DIT, presents a general strategy for further test development, and presents confirming evidence for the internal structure of the DIT. As an addendum to these papers, a table is included reviewing correlations of the DIT with various personality variables as reported in various studies.

This research is the product of many people's efforts, not only those listed as authors of these papers, but also of the scores of researchers whose work is summarized here. Their generosity in sharing their results and data have made this extensive data collection possible. Unavoidably there is much that is left out in this report, and full descriptions of the various studies are contained in the original articles which are listed in the bibliography. My collection of DIT reports, dissertations, and articles from the many contributors amounts to an eight foot stack if articles are piled on top of each other--this gives some idea of the condensation in this report. Theoretical

discussions are particularly slighted for the sake of presenting a concise summary of empirical findings. Elsewhere I have written two papers which attempt to deal with basic theoretical issues: (a) "A theoretical analysis of moral judgment development" (Rest, 1977a) describes stage characteristics and their interrelations, what the advances are of each succeeding stage, and how the higher stage provides more adequate conceptual tools for solving moral problems; (b) "The stage concept in moral judgment research" (Rest, 1977b) describes the model of development presupposed in the DIT, and how it is consistent with recent theorizing in cognitive developmental research and is consistent with what is empirically known about moral judgment development. In the near future I plan to complete a book which more adequately integrates the theoretical aspects of the DIT research with reviews of research findings, and treats a greater variety of theoretical and methodological issues.

Before going directly to the papers, a brief description of the DIT itself is in order. The subject first reads a dilemma such as the "Heinz and the Drug" dilemma used extensively in Kohlbergian research: a man's wife is dying of cancer and a druggist in the same town has a drug that might save her; however the druggist is charging an exhorbitant price for the drug and the husband can not raise the money; should he steal the drug in an attempt to save his wife? After the dilemma, the subject is presented with 12 statements which express various considerations or questions which a person might ask himself in making a decision about what one ought to do. The subject's

---

Insert Table 1 about here*

---

task is to decide which considerations or questions are crucially important

---

*Note: Tables are located at the end of each paper.

ones and which are not. More specifically, the subject is asked to rate each item on a 1 to 5 scale of "great importance" to "no importance," and then at the bottom of the page, to rank the top four items of importance, putting the choice of most important first, and so on.

In the DIT it is assumed that people at different developmental stages define the crucial issues of a moral dilemma differently. While some people see Heinz's dilemma as predominently a matter of maintaining the laws of a community like item 1, others see the dilemma more as a matter of a husband's love for his wife, like item 2. The issue statements are written as representatives of different moral judgment stages (see Rest, 1977a). The first statement is supposed to represent the Stage 4, Law and Order Orientation, the second item is supposed to represent the Stage 3, Interpersonal Concordance Orientation, and so on. Presumably if a subject has used and understands a particular stage of moral thinking, the subject will recognize that system of ideas in the item presented. For instance, a highly advanced subject presumably understands the first two statements and appreciates their relevance to this case, but also regards those statements as too narrow to serve as an adequate basis for making a decision about what Heinz ought to do. In item #5, "What values are going to be the basis for governing how people act towards each other," the advanced subject sees a much broader basis for making a decision about this case. This item incorporates concerns about social order, familial love, Heinz's own self preservation, etc. for it implicitly asks what principles can prioritize all these conflicting claims. In contrast, a less advanced subject recognizes the importance of the lower stage items but fails to comprehend the significance of the higher stage items, hence the less advanced subject gives greater importance to the lower stage items. In short, the DIT is supposed to work as a developmental measure of

moral judgment by a dual process of comprehension and preference: less developed subjects don't pick the high stage items because they don't comprehend them; more developed subjects don't pick the lower stage items because they are viewed as too simplistic.

Note however item #4 "Whether the essence of living is more encompassing than the termination of dying, socially and individually." This item was designed to be pretentious-sounding but meaningless. Such pretentious but meaningless items are sprinkled throughout the stories of the DIT--there are 6 stories in all. If a subject checks too many of these items, that questionnaire is thrown out because the subject seems to be endorsing items on the basis of their style and apparent complexity rather than on their comprehended meaning to the subject. In the instructions to the DIT, subjects are forewarned that there will be meaningless items in the test and are instructed to rate those items low. It does turn out sometimes that subjects see meaning in our meaningless items. For instance, one professor at Minnesota when he first looked at the DIT items--and in particular when he came to item #4 ("Whether the essence of living is more encompassing . . .")--offered the observation that just recently he seemed to have read a doctoral dissertation with that very same title. Now admittedly it may happen sometimes that one of our meaningless items does have <u>some</u> meaning for a particular person, but it is unlikely that the whole set of meaningless items throughout the DIT are meaningful. And so whenever a subject rates too many of the meaningless items too highly we either conclude that the subject has a different test-taking set than the one requested in the instructions, or alternatively, we infer that the subject is serving on a lot of very strange doctoral dissertation committees, but in either case, we throw the questionnaire out.

There is another circumstance in which we throw the questionnaire out: we throw out questionnaires of overly efficient subjects--that is, subjects who are so efficient with their time that they put check marks down without taking the time to read the items. Subjects who fill out the questionnaire by random checking are identified by comparing the ratings with the rankings; and if there is poor consistency between ratings and rankings, we infer that the subject is randomly checking and we discard the questionnaire. Typically about 5 to 15% of the questionnaires are invalidated either because of inconsistency between ratings and rankings or because the subject endorsed too many meaningless items.

The basic data that the DIT gives, then, are ratings and rankings to stage-keyed items to 6 stories. There are 72 items in all, some are keyed at Stage 2, some at Stage 3, Stage 4, Stage 5 and Stage 6. Therefore scores for each of these stages can be derived. Up until recently the most useful single index for the DIT has been the degree to which subjects ranked the Stage 5 and 6 items as important. I say, "up until recently" because Mark Davison will be presenting in his paper a new index that is better. But for the past 5 years we have been using an index called the "Principled" index or "P" index. This index is based on how many Stage 5 and 6 items are ranked in first, second, third or fourth place of importance. The P index is the one used in almost all of the studies presented in these papers, and it is interpreted as the relative importance that a subject gives to principled moral considerations (to Stages 5 and 6) in making moral decisions.

TABLE I -- D I T QUESTIONNAIRE

HEINZ STORY

On the left hand side of the page check one of the spaces by each question to indicate its importance.

GREAT importance	MUCH importance	SOME importance	LITTLE importance	NO importance		
—	—	—	—	—	1.	Whether a community's laws are going to be upheld.
—	—	—	—	—	2.	Isn't it only natural for a loving husband to care so much for his wife that he'd steal?
—	—	—	—	—	3.	Is Heinz willing to risk getting shot as a burglar or going to jail for the chance that stealing the drug might help?
—	—	—	—	—	4.	Whether Heinz is a professional wrestler, or has considerable influence with professional wrestlers.
—	—	—	—	—	5.	Whether Heinz is stealing for himself or doing this solely to help someone else.
—	—	—	—	—	6.	Whether the druggist's rights to his invention have to be respected.
—	—	—	—	—	7.	Whether the essence of living is more encompassing than the termination of dying, socially and individually.
—	—	—	—	—	8.	What values are going to be the basis for governing how people act towards each other.
—	—	—	—	—	9.	Whether the druggist is going to be allowed to hide behind a worthless law which only protects the rich anyhow.
—	—	—	—	—	10.	Whether the law in this case is getting in the way of the most basic claim of any member of society.
—	—	—	—	—	11.	Whether the druggist deserves to be robbed for being so greedy and cruel.
—	—	—	—	—	12.	Would stealing in such a case bring about more total good for the whole society or not.

From the list of questions above, select the four most important:

Most important _____
Second most important _____
Third most important _____
Fourth most important _____

Charting the Course of Development through Cross-sectional and Longitudinal Studies

James R. Rest

I don't think many people would be surprised to hear that there are differences in the way people define the important issues in moral dilemmas. There is a lot of variation in scores from the DIT. Generally we expect moral issues to be controversial. A question of much interest, however, is whether the differences that exist among people can be accounted for in terms of developmental differences.

The first kind of evidence regarding the developmental nature of moral judgment comes from cross-sectional data. One of the first studies collected

---
Insert Table 1 about here*
---

questionnaire data from junior high school subjects, senior highs, college undergraduates, seminarians in a liberal protestant seminary, and Ph.D. students in moral philosophy and political science. If you look at the column at the far right, you see that the relative importance given to stages 5 and 6 (the P index), is higher with the more advanced groups. Group differences in DIT scores are highly significant. About 50% of the variance in this sample of 160 subjects is accounted for by grouping the subjects into the age-educational groupings. If we assume that age and education are rough proxy variables for development, then this kind of data gives us evidence that moral judgment as measured by the DIT is developmental. There have been four other cross-sectional studies of student groups by other researchers and all of

---

*Note: Tables are located at the end of each paper.

these cross-sectional studies have found age trends in DIT scores (Sample 2 in Rest, et al., 1974; Blackner, 1975; Yussen, 1976; Martin, et al., 1977).

One advantage of cross-sectional data is its ease of collection. Researchers from all over the country have been generous in sending me reports of their findings with the DIT and in sending me their data. Thanks to their cooperation, we have a large collection of data from 136 different samples containing almost 6000 subjects. This composite sample contains data on subjects from age 13 to 74, males and females, from states from all regions of the country. Grouping the subjects by age level, Table 2 shows that the

---

Insert Table 2 about here

---

older and more advanced students have higher P% scores than the younger students. Analysis of variance across the 4 student groups gives a F ratio of 604! About 38% of the variance in scores is accounted for by grouping the students into age-educational groups.

In the non-student adult group, age is not confounded with education because an old adult can have a low education level and a young adult can have a high education level. In two studies--one by Coder (1975) and one by Dortzbach (1975)--moral judgment was more strongly correlated with education than with age. Furthermore, comparison of adult samples with student samples show that adults who ended their formal education many years ago tend to have about the same DIT scores as students currently at the corresponding level of formal education (that is, adults with a high school education have scores similar to current high school students, adults with a college education have scores similar to current college students, etc.). And so these cross-sectional studies suggest that adults in general tend to reach a plateau in moral judgment development once they leave formal schooling.

Other demographic variables have been looked at to see what variance in DIT scores they could account for. No other demographic variables are as powerful as age and education. Sex accounts for very little variance on the DIT. In 20 out of 22 studies there were no significant differences between male and females. In the 2 cases where there was a significant difference, females had higher scores. There is a little data indicating that geographical regions of the country might have some relation to moral judgment scores. I hate to offend our gracious hosts, but samples from the southern states were lower than other regions of the country. Perhaps related to this finding is the finding by Ernsberger (1976) that subjects with conservative religious affiliation have lower scores than subjects with more liberal religious affiliation. Political party affiliation and socio-economic status seem to have inconsistent relationships with moral judgment. (For a more extended discussion of this cross-sectional data, see Rest, 1976b.)

So much for cross-sectional samples relating moral judgment to demographic variables.

Although cross-sectional data is useful for suggesting many hypotheses, nevertheless longitudinal studies answer most directly the question of whether individuals change over time or whether the age-education trends of the cross-sectional studies are in large part due to cohort differences. We have a number of longitudinal studies to report. For one, we have longitudinal data on 55 junior high and senior high subjects tested in 1972, '74, and '76. Over those 4 years the average P score increased from 33 to 40 to 44. Analysis of variance produced an F-ratio of 20.1, highly significant. Furthermore, 72% of the subjects increased on the P score over that time.

A much more detailed picture comes from breaking the sample down into

---

Insert Table 3 about here

---

subgroups. One subgroup, Group J, contains junior high subjects who were 14 years old when tested in 1972 and 16 years when retested in 1974. Table 3 indicates that between '72 and '74 this group of subjects decreased significantly on Stages 2 and 3 and they increased significantly on Stage 4 and P. In the comparison between 1974 and '76, subjects had reversed their direction on Stage 4, and were shifting even more toward P. (In Table 3, the large arrows indicate statistically significant shifts and the smaller arrow indicates a non-significant trend.) Group S was composed of senior highs in 1972, aged 17. As the table shows, in the first period between '72 and '74 subjects showed significant increases in P; but in the second period, the group averages stayed the same. A third longitudinal group, Group W, was tested between '74 and '76 and the most dramatic shift was the decrease in Stage 3. The trends depicted in this table show that the P index does not completely pick up all the shifts that are occurring, in particular the shifts from Stages 2 and 3, to Stage 4. Davison will have more to say about this in his paper.

So far we have only looked at group averages rather than examining changes in individual subjects. We can look at one subject's scores, say in 1972 and in 1974, and compare each stage score at both testings. In looking at a single individual's stage scores if Stage 4 has increased, and Stage 2 has decreased, then we classify this subject as showing upward movement; if Stage 4 has increased, and P has decreased, this is an instance of downward movement. In general a change is called an upward shift if higher stages are gaining at the expense of lower stages; if higher stages are losing for the gain of lower stages, this is a downward shift. In other words the analysis of individual change is done by looking at the subject's distribution of responses. Rather than claiming that cognitive-developmental theory requires a

step by step upward movement, I propose looking for upward shifts in a subject's distribution of responses. (I have recently written a long winded paper giving the rationale for this model of development but can not go into detail here--see Rest, 1977b.)

Insert Table 4 about here

Table 4 shows a summary of individual change patterns. In Group J, between 1972 and '74, 52% of the subjects had upward changes and 12% had downward changes. The other groups are similarly represented for individual change patterns over the 2 year interval. On the average, over a 2 year interval, 53% of the subjects shifted upward and 15% of the subjects shifted downward. Over a 4 year interval, shown on the bottom of the table, upward shifts were more pronounced: on the average 66% of the subjects shifted upwards and 7% shifted downwards. These individual change patterns are similar to what Connie Holstein recently reported using Kohlberg's test (which is the most comparable study). Holstein tested 52 high school subjects over a 3 year period and found that 63% of the subjects shifted upward and 6% shifted downwards. If we take the 5 longitudinal studies all together that have used Kohlberg's test, the longitudinal results using the DIT are fairly similar to longitudinal results from Kohlberg's test, as Table 5 shows. (Table 5 includes only those longitudinal studies which have not changed the scoring system while scoring the data.)

Insert Table 5 about here

In Rest, 1975b, I emphasized the benefits of going to college because the most dramatic changes between 1972 and 1974 were in those subjects leaving high school and going to college. However, between 1974 and 1976 the most

dramatic changes were in subjects leaving high school and going to college or going to work. Therefore, I have to modify what I said in my 1975 report. Now it appears that the most helpful thing you can do for a person's moral judgment isn't necessarily getting them into college, but is just to get them out of high school. At this point it is unclear whether there has been a shift in just 2 years in the type of person staying out of college, or whether the original findings which found so much change associated with going to college were just a fluke. In any case, much more detailed analysis of specific experiences and how they are linked to changes in moral thinking is necessary. We will want to know more specifically what it is about college experiences and work experiences that foster development and what it is about most adult experience that seems to stifle further development.

In conclusion, both cross-sectional studies and longitudinal studies provide evidence of developmental trends in the way people define moral issues. The two types of studies along with the replications provide validating support for the Defining Issues Test. If we are given demographic data on a sample of subjects, we have a pretty good idea of what their group averages will be on the Defining Issues Test.

Table 1

Group Differences on the D.I.T. Indices

Student Group		STAGE 2	3	4	5 and 6 (P)
Junior High	$\bar{X}$ *	11.6	20.5	35.2	32.7
n=40	SD	(7.3)	(9.8)	(11.8)	(14.1)
Senior High	$\bar{X}$	9.6	22.3	30.7	37.4
n=40	SD	(6.7)	(10.1)	(11.4)	(15.4)
College	$\bar{X}$	5.5	14.6	24.9	54.9
n=40	SD	(4.8)	(7.8)	(10.7)	(13.6)
Grad	$\bar{X}$	3.5	13.0	18.4	65.1
n=40	SD	(4.5)	(10.7)	(7.9)	(11.7)
a) Seminarians (n=25)	$\bar{X}$	4.7	15.5	17.9	61.9
b) Political Science and Philosophy Majors (n=15)	$\bar{X}$	2.2	8.8	18.8	70.3

* $\bar{X}$ is the average percentage of ranks (weighted 4 for 1st rank, 3 for 2nd rank, 2 for 3rd rank, 1 for 4th rank) given to the "issues" of each stage, respectively. One way analysis of variance between groups on the P score produced an F = 48.5 (F at the .01 level of significance $\geq$ 3.95).

Table 2

Large Group Means and Standard Deviations

Group	n	Average	Estimated S.D.	Range of Most Sample Means
Junior high	1322	21.9	8.5	20.0 - 26.7
Senior high	581	31.8	13.5	26.7 - 36.7
College	2479	42.3	13.2	36.7 - 46.7
Grads	183	53.3	10.9	53.3 - 60.0
Adults	1149	40.0	16.7	36.6 - 50.0
	5714			

Table 3

Subgroup Stages Changes Over Two Years

Group	Ages	n	STAGE CHANGE 2	3	4	P
Ja + b	14 - 16	50	↓↓	↓	↑	↑
Ja	16 - 18	31	—	—	↓	↑
Sa + b	17 - 19	38	↓	↓	↓	↑
Sa	19 - 21	23	—	—	—	—
W	18 - 20	21	↓	↓	↑	↑

Table 4

Individual Pattern Changes Within Subgroups

Two Year Interval

Group	Ages	n	Percent Ss Moving Up	Percent Ss Moving Down
Ja + b	14 - 16	50	52	12
Ja	16 - 18	31	53	16
Sa + b	17 - 19	38	66	5
Sa	19 - 21	23	39	39
W	18 - 20	21	48	14
			Average up = 53%	Average down = 15%

Four Year Interval

Ja	14 - 18	31	69	6
Sa	17 - 21	23	63	8
			Average up = 66%	Average down = 7%

TABLE 5

Upward and Downward Movement in

Kohlbergian Longitudinal Studies

Study	Time interval between testings	Age of Subjects	Number of cases	Percent moving Up	Percent moving Down	Ratio of Up to Down movement
Kramer, 1968	3 years	high school	24	42	17	2.5 to 1
	3 years	college	19	21	16	1.3 to 1
Blatt & Kohlberg, 1973	1 year*	12-13 years	10	20	60	1 to 3
Holstein, 1976	3 years	high school	52	63	6	10.5 to 1
	3 years	adults	97	29	18	1.6 to 1
Kuhn, 1976	1/2 year	5-8 years	100	44	24	1.8 to 1
	1 year	5-8 years	50	64	10	6.4 to 1
White, 1977	1 year	8-17 years	242	47	21	2.2 to 1
	2 years	8-17 years	86	87	8	10.9 to 1
			Average:	50.2	17.7	

*Experimental subjects in Study 1; comparison between posttest and one year followup.

Cognitive and Political Attitude

Correlates of the Defining Issues Test

James L. Carroll

Arizona State University

One facet of construct validation is the examination of correlational relationships of the Defining Issues Test (DIT) with other psychological measures. As Mischel (1976) and Alston (1971) have pointed out, age trends do not completely validate a measure of moral judgment. More is implied in the cognitive developmental notion of moral judgment than just that there are age trends. Exploring some of those additional implications, this paper summarizes data from about 40 studies on the correlations of the DIT with cognitive developmental measures, IQ and academic achievement measures, and political attitudes.

One question of major theoretical interest is whether changes in moral judgment reflect changes in a person's capacity for higher stage thinking. Merely showing that people's thinking does change over time does not demonstrate that their capacity has changed over time. It could be (as a logical possibility) that young children have the capacity to think in the higher stages but that they choose not to do so; cross sectional or longitudinal change on measures of moral judgment might only represent changes in preference for certain kinds of thinking rather than increasing cognitive capacity.

In order to examine cognitive capacity, a different kind of measure is required than either the DIT or Kohlberg's test. It is for this purpose that the test of moral comprehension was devised.

Table 1 about here

Table 1 gives an example of the moral comprehension test devised by Rest. The subject is presented with a paragraph and then below that paragraph are four statements. The subject's task is to select from the four statements the one which best reflects the meaning of the paragraph. Note that the subject is not asked whether or not he agrees with the paragraph or the statements; the subject's job is to decide which statement comes closest in meaning to the paragraph. It is presumed that those subjects who correctly match the statement with the paragraph understand the concept illustrated in the paragraph. Eleven paragraphs are used in the Moral Comprehension Test, illustrating concepts usch as social contract, legitimate authority, autonomous thinking based on principle. Comprehension scores range from 11 to 0, depending on the number of correct matches that the subject makes. Moral philosophy and political science doctoral students (our "expert" group) consistently choose the alternatives which are keyed as correct, and other groups of subjects get lower scores.

Table 2 about here

Table 2 shows the correlations of moral comprehension with the DIT. In all but one comparison, the correlations were significant. In heterogeneous groups the correlations go into the .60's, but in more homogeneous groups the correlations are lower. In the sample of 73 9th graders, the correlation between the DIT and comprehension was .50 after age was controlled and IQ, socio-economic class, and sex were statistically partialled out. In the longitudinal study reported by Rest , comprehension increased significantly just as did the DIT -- as comprehension went up, so did the DIT: Comprehension scores went from an average of 5.1 to 7.3 between 1972

and 1976, $t = 6.8$, $p < .001$. Of the subjects who showed upward movement across DIT testings, 81% also increased in comprehension. In summary, moral comprehension -- a test of cognitive capacity -- is fairly well related to DIT scores.

A second question of interest is how the DIT is related to Kohlberg's test of moral judgment. Since the DIT is derived from Kohlberg's approach and his characterizations of stages, we would theoretically expect the two measures to be related. Also it is of practical interest to know whether the DIT is an equivalent test to Kohlberg's.

Table 3 about here

Table 3 indicates that correlations in several studies ran as high as the .70's. The highest correlations occurred with heterogeneous subject groups. In more homogeneous groups, the correlations were considerably lower. In general, the various forms of Kohlberg's test (or tests derived from Kohlberg's test) seem to be related to the DIT, but not closely enough to regard the measures as equivalent.

Alozie has done the most intensive comparison of the DIT and Kohlberg's test. At first Alozie determined how much difference between the DIT and Kohlberg's test could be attributed to certain design differences in the tests. For one, the DIT uses a somewhat different set of hypothetical dilemmas than Kohlberg's test. Both Kohlberg's test and the DIT use the familiar Heinz and the drug dilemma, but the measures match on only half of the dilemmas. Alozie's findings indicate that the correlations between the DIT and Kohlberg's test weren't much higher for the matching dilemmas than for the dilemmas that differ.

Secondly, Alozie examined the effect of some differences in stage characterizations. Rest has defined the 6 stages somewhat differently than

Kohlberg, and therefore some DIT items are stage keyed differently than they would be within Kohlberg's present system. Accordingly, Alozie went over DIT items and rekeyed items to conform more closely with Kohlberg's stage definitions. Alozie, however, found that changing the stage definitions of some DIT items did not improve the correlations with Kohlberg's test, but in fact lowered the correlations a little bit.

Thirdly, Alozie checked out the effect of each test using a different kind of index: Kohlberg's test uses stage typing and the moral maturity score; the DIT uses the P score. Alozie asked if computing similar indices for both tests would give a higher correlation between the tests.

Table 4 about here

Table 4 shows how the various indices correlate. We see that the P index for the DIT and the moral maturity score from Kohlberg's moral judgment interview give the highest correlation, .75 -- therefore equating the two tests on type of index used does not improve the correlation between them. (It should be noted at this point that Davison's new index was not available at the time of this study, and we don't know how that compares as yet.)

Alozie concluded that these design differences in the tests don't account for much of the difference between them. Alozie then examined the most striking difference between the two tests: the tendency for the DIT to score higher than Kohlberg's test. In fact, because the DIT credits subjects with much higher thinking than Kohlberg's test some researchers have doubted that the DIT could be measuring moral judgment at all. Kohlberg has recently said that principled moral thinking does not occur until adulthood, and even then it is a rarity. However the DIT has even junior high students attributing some importance to principled thinking, and on the DIT, principled thinking is

not a rarity at all.

Alozie carefully examined scores from Kohlberg's test and scores from the DIT for each subject. He found that on the average, Kohlberg scores were almost two stages below DIT scores (Table 5). For instance, Alozie examined Kohlberg scores on the subject's responses to the Heinz dilemma. Alozie found that subjects did not tend to pick DIT items at the same stage as on the Kohlberg test, but tended to pick items at stages above.

Table 5 about here

Alozie found that in 77% of the comparisons, subjects chose DIT items that were at stages higher than the stage scores on the Kohlberg interview.

This discrepancy between the DIT and the Kohlberg test has parallels in Piagetian research and in other areas of social cognition. The DIT is essentially a recognition task whereas Kohlberg's test requires subjects to rebally produce and justify an answer. With such a difference in response mode it is not unusual in cognitive developmental research to find that the recognition task (the DIT) is easier for subjects than the production task (the Kohlberg test), and to find that the DIT credits subjects with higher level thinking earlier than the Kohlberg test.

In conclusion, the DIT and the Kohlberg tests are not interchangeable, although the correlations between them are consistently positive and usually significant.

The DIT is correlated with other measures of cognitive development which are not distinctively measures of moral judgment.

Table 6 about here

As Table 6 shows, these relationships are generally positive, but not as high as the DIT's correlations with measures of moral cognition. In addition, there is data on about 50 correlations of the DIT with IQ measures and measures of academic achievement and grades. Space does not allow detailed discussion of these correlations, but they are generally positive, significant, and in the .20's and .40's range. Again this is evidence of a relationship of cognitive capacity with the DIT, however the magnitude of these correlations with IQ, etc. is less than that of the DIT with measures of moral cognitive development.

A third question that correlational data can help address is whether the DIT is measuring a value-free, purely intellective skill, or whether the DIT has anything to do with values. One measure is the "Law and Order" test, devised by Rest.

Table 7 about here

The Law and Order test asks subjects to take a stance on current controversial issues of public policy. Several studies have reported significant correlations of the Law and Order Test with the DIT (Table 8).

Table 8 about here

Of the ten correlations, 9 report a significant correlation: the higher the DIT score, the less tendency for a subject to endorse value positions that attribute almost limitless power to authorities or that value maintenance of social institutions at high costs to individual welfare and freedom. In Rest's longitudinal study, as DIT scores went up, Law and Order scores went

down (t = 4.2; p < .001). Of subjects who increased on the DIT, 84% decreased on Law and Order.

Correlations of the DIT with other measures of political attitudes are generally lower and inconsistent. For instance, correlations with measures of Political Efficacy, Political Interest, Political Tolerance, Hogan's Survey of Ethical Attitudes, Rokeach's Dogmatism, etc. are inconsistent and usually not very large. Table 9 provides an overview and references for many of these measures.

Table 9 about here

Other questions regarding the relationship of the DIT to values, attitudes, and behavioral acts are further discussed in Jacobs and McColgan.

Table 1

Sample Comprehension Item

If Heinz steals, he is breaking his agreements with other members of society. In most countries mean have agreed not to steal because they see that not stealing is better for each one of them. Heinz himself would have to admit that a law against stealing is a good law to have. And so if Heinz wants to have laws that he and other people think are good to have, he should abide by them.

____ Good	____ Fair	____ Poor	a.)	Men consent to laws because they recognize that in the long run laws benefit each member of society. One's obligation to obey the law comes from this recognition.
____ Good	____ Fair	____ Poor	b.)	Heinz should not steal because if he does, people will think he has broken his agreements with them, and they would regard him as untrustworty.
____ Good	____ Fair	____ Poor	c.)	Once the law is set, no one is right in breaking it. No matter what good intentions a person may have, if he breaks the law, he's in the wrong.
____ Good	____ Fair	____ Poor	d.)	Heinz has a duty to obey the law because he helped to make the law. If he breaks his agreements, he will be setting an example that could lead to everyone's breaking the law.

Rank the statements from best interpretation of the paragraph (1), to worst (4).

1. ____, 2. ____, 3. ____, 4. ____

Table 2

Correlations of DIT with Moral Comprehension

Study	Sample	r
Alozie, 1976	91 junior highs and college students	.68
	(37 college students only	.46)
Rest, et al., 1974	67 Ss, ages 14 to adulthood	.67
	160 Ss, junior high to graduate school Ss	.62
	(9th graders only, n=73	.58)
Coder, 1975	87 adults (age 24 to 49)	.49
Rest, 1975	88 Ss, age 17-20	.42
Masanz, 1975	34 high school girls	.37
McColgan, 1975	52 junior high predelinquents and controls	.34
	29 delinquents	.19

Table 3

Correlations of DIT with Kohlbergian Tests of Moral Judgment

Study	Sample	Measure	r
Alozie, 1976	91 Ss, combined group (37 college Ss (52 junior highs	Kohlberg's 1972 Issue Scoring	.75 .28) .21)
Gibbs & Fedoruk, 1975	41 college students	Gibbs' scale of principled moral thinking	.70
Rest, et al., 1974	47 students and adults	Kohlberg's 1958 Scoring System	.68
Froming & McColgan, 1977	81 adolescent boys & 163 college Ss	Kohlberg Scoring System (recent version)	.65
McColgan, 1975	29 delinquents	Kohlberg's 1972 Issue Scoring	.37
Carroll & Rest, 1977	88 Ss, 18-22 yr old Ss	Written and modified version of Kohlberg's 1972 Scoring System	.34
	55 Ss, 20-24 yrs.	Kohlberg's 1972 Scoring System	.41

TABLE 4

Correlations of the Kohlberg Test with the DIT on Three Indices

		Kohlberg Test Index		
		P	MMS[a]	Stage Typing
DIT	P	.45	.75	.64
	MMS	.42	.72	.66
	Stage typing	.33	.61	.55

[a] A Moral Maturity Score

From Alozie, 1976

TABLE 5

Percent Subjects Scoring Higher, Same,
or Lower on DIT than on the Kohlberg Test

Story	Higher on DIT	Same on DIT and Kohlberg	Higher on Kohlberg
Heinz	64	27	9
Doctor	71	14	7
Prisoner	87	6	8
Total	77	16	8

n = 91

From Alozie, 1976

Table 6

Correlations of DIT with Other Cognitive Developmental Measures

Study	Sample	Measure	r
Briskin, 1975	32 college Ss	Leveling-Sharpening	.49
Meyer, 1975	40 male college Ss	Perry's Intellectual & Ethical Development	.45
Panowitsch, 1976	82 college Ss	Cornell Critical Thinking Test	.41
Cauble, 1976	90 college Ss	Piaget's Formal Operations	.40
McColgan, 1975	29 delinquents	Piaget's Golden Rule Task	.30
	52 predelinquents	Piaget's Golden Rule Task	.31
	52 predelinquents	Chandler's Cartoon Role-Taking	-.18
Copa, 1975	127 college women in home-ec. classes	Harvey's Conceptual Systems	.13

Table 7

Sample Law & Order Item

Under present laws it is possible for someone to escape punishment on the grounds of legal technicalities even though the person may have confessed to performing the crime. Are you in favor of a tougher policy for treating criminals?

_____strongly agree with tougher policy
_____mildly agree
_____mixed agreement and disagreement
_____mildly disagree
_____strongly disagree

If a person is against a war that his country is engaged in, is it right to do things which disrupt the war effort (like destroying government records, disrupting government buildings, demonstrating at army installations, etc.)? (Check one)

_____never right
_____sometimes right
_____right most of the time
_____don't know

Table 8

Correlations of DIT with Law & Order Attitude Test

Study	Sample	r
Rest, et al., 1974	160 students--jr. high to grad Ss	-.60
	(73 junior highs	-.23)
	65 students--jr. high to adults	-.48
Rest, 1975	88 Ss, 16 to 20 yrs.	-.52
Coder, 1975	87 adults	-.49
Rest, 1976	55 Ss, 18 to 22 yrs.	-.47
Panowitsch, 1975	44 college Ss	-.27
Masanz, 1975	34 high school girls	-.19

Table 9

Attitude Correlates of DIT

Study and sample	Attitude test and Variable	Correlatio
Allen & Kickbusch, 1976	Dean alienation scale:	
	- Powerlessness	.04
410 9th graders	- Normlessness	.18
	- Social isolation	.06
	Political efficacy	.16
	Political trust	-.03
	Salience (interest in political courses)	-.08
	Political aspiration	.01
Coder, 1974	Radical-conservatism scale	.13
58 adults		
Gallia, 1976	Rokeach dogmatism	-.75
10 college science majors		
10 college humanities majors	Rokeach dogmatism	-.03
Gutkin & Suls, 1976	Hogan's Survey of Ethical Attitudes	-.27
68 college Ss		
Morrison, Toews, & Rest	Portune's Attitudes Towards Police	-.13
1973	Political tolerance	-.08
71 junior highs	Political efficacy (I)	-.13
	Participation	.12
	Political Trust	.07
	Political efficacy (II)	.17
	Civic tolerance	.10
	Political interest	.11
	Rokeach Dogmatism	-.28
Rest, 1975	Political tolerance	.21
88 high school Ss		
Rest, Ahlgren & Mackey, 1973	Portune Attitudes Towards Police	-.37
61 junior highs	Political tolerance	.12
	Political efficacy	-.17
	Copo-Polo Scale	
	- view police as helpful	.23
	- angry feelings towards police	.15
	- attribute concern to police	-.10
Rest & Feldman, 1973	Political efficacy	.04
160 students	Political activity	-.19
	Political positiveness	-.34

MORAL JUDGMENT INTERVENTION STUDIES USING THE DEFINING ISSUES TEST

Jeanette A. Lawrence
University of Minnesota

Intervention studies provide information related to a) the DIT and its construct validity, b) the nature of moral judgment, and c) the design and execution of a developmentally informative moral judgment intervention study.

Moral development interventions usually follow one of a few well-used models: moral dilemma discussions, modeling of higher levels of moral reasoning, or stimulation of general psychological (including moral) growth. They had traditionally used the Kohlberg interview and scoring techniques to index changes in their subjects' levels of moral thought. The field is in need of new directions for innovative, theory-related, rigorous research. What has been achieved has not been systematically reviewed prior to Lockwood (1977) and this present review. Lockwood provided a much needed critique of interventions in the Values Clarification and Kohlberg traditions. The present review examines fourteen studies which use the DIT to index moral judgment development following deliberate attempts to raise subjects' levels of general psychological or moral development. The studies are the work of a variety of researchers with divergent objectives and interests. They have varied sample characteristics and experimental designs. Eight of the studies produced significant upward movement on the DIT. This critique of the body of research found in the studies provides independent support for the ability of the DIT to measure upward movement in moral thought. The DIT does index upward change in moral judgment scores after educational treatments specifically designed to promote development.

Table 1 presents a summary overview of the studies; their sample sizes and characteristics, design and focus, instrumentation and testing; and evidence of change on the DIT. The studies have been grouped for convenience according to two criteria; duration, and type of treatment experience provided in the educational program. These programs were in the areas of social studies, psychological development, or moral development education.

TABLE 1 ABOUT HERE

While no blueprint for the successful moral judgment intervention emerges, there are several indicators of what will contribute to confidence in the experiment's claim to have facilitated change: the duration, design and characteristics of the experimental program. Walker's study is an example of the type that might well be dropped from moral education programs, i.e., the brief, one-lesson treatment. This kind of study assumes that structural and conceptual change can be produced instantly. Structural change involves the reorganization of the person's way of thinking about moral issues, and that pre- and posttestings, and for effects of time and stimulus variables on the internal processes of development within the subject.

Moral judgment treatments are not concerned with overt skill acquisition, but with changes in the subject's mode of thinking about moral issues. In order to infer that the treatment stimulated such change the experimenter must provide: adequate control groups, control of extraneous variables within the treatment group; description and observation of variables assumed to be developmentally critical; some monitoring of the assumed mechanisms of structural change; and developmentally appropriate testing.

While all the experimenters claimed some sort of developmental significance for their studies, two did not use control groups at all (Rest, Algren, Mackey, 1972; and Erickson, et. al., 1975), and one only minimally (Sprinthall and Bernier, 1977). Several used levels of control (Hurt, 1975, Panowitsch, 1975, Whiteley, 1976), and others alternative moral treatments (Coder, 1975, Piwko, 1975, Siegal, 1974).

Studies varied in the degree of control exercised over treatment variables and their relation to developmental or skill-related dependent variables. The reports of multitreatment studies make it difficult to know which of the numerous treatment factors actually influenced development. They ranged widely across such things as: teaching moral development theory, journal-keeping, empathy training, residential and community experience, and often seemed to be thrown in as a pot pourri. Erickson et. al., Hurt, and Sprinthall and Bernier tried to monitor the effects of one skill, i.e., empathy training. Inclusion of auxiliary testing is an advance towards identification of effective moral judgment treatment variables.

Added to sampling and control constraints on the experimenter's claims, are statistical analyses and testing problems. Without randomization, the studies were dependent on pre- and posttest comparisons between experimental and control groups for some kind of equalization. Kickbusch (1976) and Siegal (1974) used analysis of covariance, with the pretest as the covariate, while the others mostly compared groups on pretest measures. Only Coder, 1975, Morrison, Toews and Rest, 1975 and Panowitsch, 1975 compared experimental and control groups on posttests. This omission of data in the other reports implies results were reported selectively. As well as indicating

the need for the kind of empirical rigor which deals with these issues, the group of studies give clear indication of the need for theoretical direction for intervention studies.

The illustrative study, Panowitsch's, shows how attention to theoretical concepts of "moral" and "judgmental" -- can be operationalized, and can complete the cyclical relationship between theoretical concepts, theoretically designed interventions, and theoretically valid instrument, i.e., the DIT whose construct validity is supported by the independent studies. Panowitsch illustrates the value of a moral judgment intervention using moral orientation and judgmental problem-solving. The DIT and Cornell Critical Thinking Test (CCTT) were given to 4 types of college classes--Ethics, Logic, World Religion, and Art (the latter two acting as controls). One specific aim was to test the sensitivity of the DIT in differentiating gains in moral critical thinking (in the Ethics Groups), from general critical thinking (in the Logic Groups), and general values content (in the World Religion and Art Groups). It was expected that if the DIT tested only conceptual change, the ethics and logic groups would not differ in their posttest gains on the two instruments. If the DIT were simply a test of general values orientation, then the ethics and the world religion groups would not differ on the DIT. But if the DIT were a sensitive test of moral judgment, then the Ethics groups would show significantly higher gains than the other two groups.

The Ethics classes' treatment also provided opportunity to examine the content of a moral judgment intervention that used a treatment built on the concept that moral judgment-making involves solving moral problems in relation both to their content, and the problem-solver's judgmental processes. The Ethics courses had two components addressing this idea: a) reading and

understanding the thought of moral philosophers, e.g., Kant, Neitzsche, and b) the application of methods of moral problem-solving to contemporary issues, e.g., suicide, abortion, civil disobedience. Thus students were given experience in studying moral concepts, and in making their own moral judgments. In contrast, the Logic students were given training in formal logical problem-solving. World religion and Art students had experience with broad values issues.

Results confirmed hypotheses of the interrelationship between moral judgment theory, the ethics course, and the DIT, P. index. The combined ethics groups showed significant gains on the DIT from pre- to posttest $t(72) = 3.21$ $p \leq .002$, while the logic and world religion and art classes showed no significant gains. Logic groups showed the only significant gains on the CCTT over the same length of time $t(22) = 2.19$, $p \leq .040$. The logic and ethics groups did not differ on the DIT pretest but were significantly different on the posttest $t = 2.16$, $p \leq .034$, (see Figure 1). Comparisons by the Scheffe test showed that the Religion and Art classes differed from the Ethics and Logic classes, but did not differ significantly from each other ($p \leq .05$).

FIGURE 1 ABOUT HERE

Panowitsch gave a five months follow-up test on the DIT to the students taking the Ethics and Logic courses in the spring quarter. Figure 2 shows the data results for these subsamples.

FIGURE 2 ABOUT HERE

The changes for these Ethics subjects were significant for both pre- to posttest ($p \leq .003$), and pretest to follow-up ($P \leq .015$). From posttest to follow-up they did not change significantly ($p \leq .870$) but showed retention, consistent with theory. The logic subject showed no significant gains from pre- to posttest, or from either pretest or posttest to follow-up. This study illustrates the value of: a) designing a moral education program in accord with the theoretical concepts of moral and judgmental components of moral thought; b) the DIT as a useful index of moral judgment for educational interventions c) the use of a large sample, i.e., 152 students and d) the addition of follow-up testing in a developmentally oriented study.

In that Panowitsch's subjects were not assigned to treatments randomly, but were in self-selected intact classes, there is a possible confounding of student interest and course content. To control for possible differential selection effects, spring Ethics and Logic classes completed questionnaires asking students for their reasons for selecting either the ethics or the logic course. Reasons given for selecting the two courses were similar, and this was supported by the absence of pretest differences between the two groups on the DIT and CCTT.

Apart from concerns for rigor and economical, efficacious and appropriate use of program resources, two important questions are raised by such a review as this: "What causes an experimenter to adopt a developmental perspective for his study?" And, "Why choose moral judgment development as the dependent variable?"

Why adopt a development perspective?

Invariably moral and psychological education programs are claimed to have developmental significance. Yet it seems that experimenters do not weigh

the values and costs of adopting a developmental framework for their programs. An obvious advantage is the purchase of contextual meaning for the results, both in relation to the subject's life-span, and in plugging into a body of research on developmental trends. It is easy to tread the well-worn pathway, gain contextual meaning, and fail to identify developmentally influential variables. Thus unmonitored treatments, or those invoking unexplained mechanisms, may proliferate reports, but they do not advance knowledge of moral judgment development.

Yet there are costs to exchanging isolated interpretation for the fellowship of the developmentally oriented. There are some major problems inherent in developmental research. For instance, how can gain scores on a particular index be interpreted in terms of actual growth in moral reasoning? There is no firm ground for assuming that scores on our rather crude measures sensitively mirror changes in subject's mental operations. Mark Davison's (1977) paper addresses the problems and possibilities of indexing moral judgment scores, and see Rest (1977b).

Even if moral judgment levels could be accurately diagnosed and measured, interpretation problems are compounded by situational, personal, and interactive effects, and the test-taking factors that can intrude between the person's moral reasoning and test scores. Change scores may simply reflect either the subject's increased understanding of the test task, or good feelings about the experience. Paradoxes abound in the test-taking complexities of such distinctions as competence and performance; structure and content; and hypothetical test and real-life moral problems.

Why moral judgment as the dependent variable?

If the developmental perspective is suitable, or chosen, for the study, why moral judgment development? Some of the enthusiasm may be due to the DIT's ease of administration and analysis. It can be given as a paper and pencil test in a group situation, and is computer analyzed. Another factor is the scarcity of developmental measures. The DPE model tries to build up a composite picture of general conceptual growth, and includes the DIT in the dependent variables. Yet it seems there ought to be a rationale for using a measure which is specifically "moral" and "judgmental" in content. Results show that the DIT indexes a particular sort of change in general psychological development programs, and furthermore, correlational studies comparing it with other developmental measures may confirm that it is as useful as some cognitive and ego development indexes of general conceptual change. But its purpose is to index one particular aspect of human thought and morality: the moral judgment. Panowitsch's study shows its ability to do this.

This review has shown the relative uncertainty with which moral educators can base their programming on research. Lockwood's review of Kohlbergian studies (1977) also questions the strength of current models of moral development interventions. Both the Siegal and the Coder results raise serious questions about the dilemma-cum-discussion method for moral education programs. Coder's lecture treatment was as effective as the standard Kohlbergian approach. Siegal's adaptation of the discussion method was no more successful than the attempt to influence upward movement by developing subjects' moral reasoning skills. Little is known about effective environmental variables, especially in multi-treatment studies. Still less is known about organismic processes of change. There is need for more experimental investigation of the twin concepts of moral criteria and judgmental skills.

Directions for further research

How can moral judgment intervention studies be made more developmentally informative? Apart from adherence to sampling, control and testing standards, objectives need to be made explicit in relation to moral judgment theory, and more accurate diagnostic measures need to be devised. If the DIT is used, it should be for its theoretical assumptions, and their suitability to the independent variables. In-put variables believed to be developmentally influential need to be monitored and measured: a) intraexperimentally, b) in relation to the developmental dependent variables, and c) for cumulative, generalizable, and continuing, i.e., developmental, outcomes over time, in follow-up tests. If listening skills, discussion participation, or log-keeping are assumed to stimulate growth, then the experimenter needs to report, for example, how much a subject took part in discussion, and how his performance correlates with DIT scores, and how it compares with that of the non-participant. This is one way of linking critical variables to developmental outcomes.

Finally the study should be placed within a research program of replication and refinement. A number of the studies reported here were either dissertation studies or the experimenter's first attempt at operationalizing some hunches, or a particular program. This factor in itself places certain constraints on what can be achieved. In the first run, the experimenter often has limited control over developmental variables, curriculum materials, teacher training or the sometimes strange climate of the classroom, as well as experimental and design factors. The first attempt should not ordinarily be expected to validate procedures for theoretical presuppositions, but rather, to link present research to past findings, and to identify variables of developmental interest.

Placed within an ongoing program, the findings of the study can feed back into the next wave of research. Transition studies should be used to sort out effective variables from "multi-treatment interference" (Bracht and Glass, 1968). In summary three specific suggestions emerge from this review: interventions should be tailored to fit moral judgment theoretical concepts, the measurement instrument should be a good fit for the treatment, and future research should be focused on moral and judgmental aspects of moral judgment programs.

TABLE I. OVERVIEW OF INTERVENTION STUDIES, USING THE DIT.

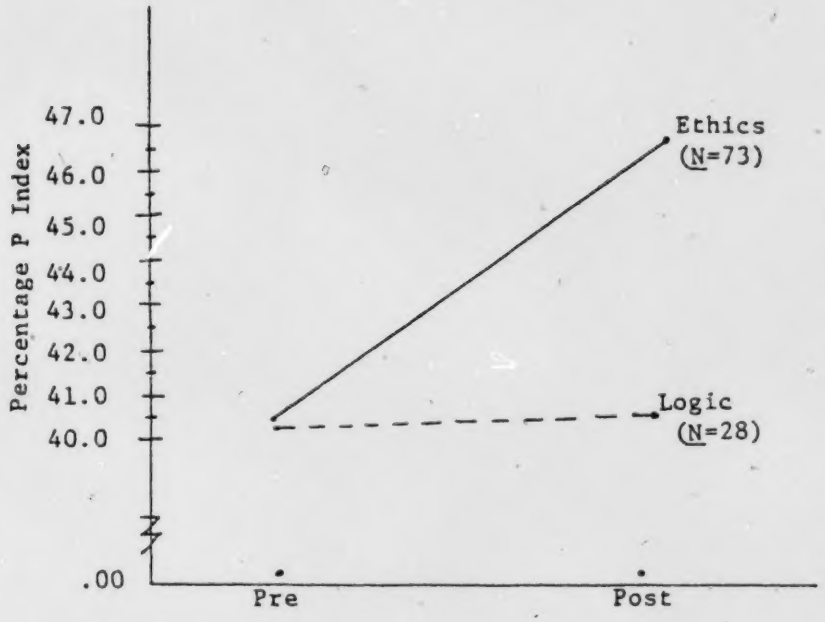
Study	Sample	Type of Intervention	Duration	Design	Change on DIT	Characteristics
SHORT-TERM STUDY:						
1. Walker, 1974	8th grade N=70	narrative modeling of pro, con reasoning (+1), (-1), own stage	one-lesson treatment	3 exp. groups (+1), (-1), (0)	Es: no change	assumes immediate, easy change and stage typing
SOCIAL STUDIES PROGRAMS:						
2. Mpls. Police Report, 1972	Jr. high N=61	soc. studies instructional unit to change attitudes to police	12 weeks	exp. groups no control	E: no change on P, decrease in use of St.2, t=-2.17, p<.05	attitude change study, with developmental measure, no control group
3. Morrison, Toews, Rest, 1975	Jr. high N=103	"Youth, Law and Morality" Program: social & civics studies, to develop moral reasoning	4 months	exp. group control groups (follow-up test six weeks)	Es: no change Cs: no change	multivariable treatment use of follow-up test
PSYCHOLOGICAL DEVELOPMENT PROGRAMS:						
4. Erickson, Colby, Libbey, Lohman, 1975	Jr. high N=20	DPE Curriculum for personal development, within school classes	semester	exp. group, no control	E: t (19) = 2.27, $p \leq .02$	auxiliary empathy scale taught Kohlberg stages (no control group)
5. Hurt, 1975	College N=54	counselling, empathy training, in educational psychology course	quarter	exp. group, (E) active con. (C1) inactive con(C2)	E: $t(14)=1.94$, $p\leq.037$, C1: $t(18) = 1.80$, $p\leq.045$, C2: NS	treatments randomly assigned to intact groups, aux.testing, 13 inconsistents in E. group
6. Sprinthall Bernier, 1977	in-service teachers N=18	intensive workshop in personal, professional development, plus seminar while teaching	6 weeks workshop, + qrter.	exp. group	E: $t(17)=2.91$, $p<.01$	taught theory, skill index aux. testing, comparison group to relation to E
7. Balfour, 1974	Sr.high N=84	"Humanities Outreach Course": community experiences, seminar	semester	3 exp. groups, in 3 schools, 1 control	Comb.E: $t(53) = 2.01$, p .05; West.E: $t(11)=$ 1.83, $p\leq.046$ C: no change	multivariate treatment, transfer to DIT D.V.
8. Kickbusch, Allen, 1976	9th grade N=117	"Confluent Education" for integration of affective, cognitive aspects of curriculum teaching style, unit on morality	8 months with semest. in moral educ. unit	exp. groups (E) Control (C1) Control (C2)	E: no change C1: no change C2: no change	multiplicity of treatment variables and testing, inconsistency norm on DIT

over

Study	Sample	Type of Intervention	Duration	Design	Change on DIT	Characteristics
9. Sierra Project, Whiteley, Nelsen, 1976	College freshmen N=77	Residential, community living program to integrate formal and informal education	8 months, with moral education courses	exp. group, placebo con(C1) control (C2)	E: $t(34)=2.37$, $p\leq.024$, C1: $t(26) = 1.6$, $p<.12$(NS), C2: no change	multiple experiences, course module on moral dev. theory. Ran ass. of volunteers to E. and C1, ran sample C2
MORAL EDUCATION PROGRAMS:						
10. Troth, 1974	College N=42	"Values Course": to integrated personal values and behavior	semester	exp. and control groups	E: no change C: no change	pretest (1/2 DIT), posttest (whole), new program
11. Siegal, 1974	8th, 9th 10th grade N=358	Treatments: I. Kohlberg, II. Meux, III. Aver, IV. Inquiry/control	semester	3 exp. groups, 1 control	E: no change C: no change	large scale, confounding of treatments, materials, and teacher training
12. Piwko, 1975	College freshmen N=36	"Moral Development Workshop": moral values, commitment, Human development course (C1) no treatment (C2)	quarter	ext. groups, 2 controls	E:$F(1,33)=6.89, p<.05$, Cs: no change	course focused on exploration of moral issues
13. Coder, 1975	adults: church members N=87	A. Dilemma and discussion, B. Lecture, no discussion C. Other seminars	6 weeks, (x 2 hrs.)	2 exp. groups, (A and B) 1 Control (C)	(A+B) differs from C on Posttest:$F=5.69, p<.005$ NS diffs. between A & B	discussion lecture, (not significantly different) delay in return of posttests
ILLUSTRATIVE STUDY:						
14. Panowitsch, 1975	College N=152	E: Ethic class, L: Logic class, World religions class (C1) Art class (C2) comparison of moral judgement, critical thinking, and general values treatments	quarter + 5 months follow-up for 2 sub-samples	2 exp. groups, 2 controls (follow-up 8 months)	Comb. E: pre-post DIT $t(72)=3.21$, $p\leq.002$, Comb. L: pre-post CCTT, $t(22) = -2.91$, $p\leq.040$, Controls NS, E vs L: pre-NS, post-$t = 2.16$, $p \leq .034$ E: post- to follow-up, NS.	large sample, levels of control, use of follow-up, operationalized "moral", "judgmental" content, intact classes.

Figure 1

Matched Pairs t Test Pre to Post D.I.T. Percentage P Index for Total Ethics and Logic Groups After the Consistency Check*



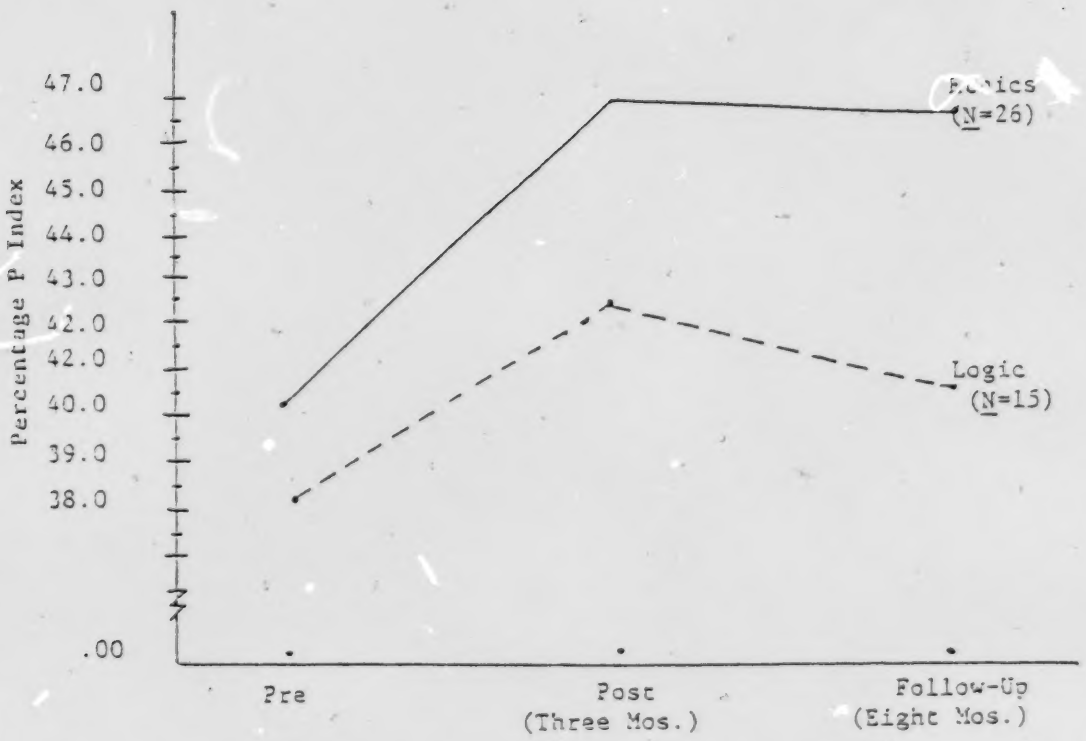
*Only protocols which pass the criteria of the consistency check were used in the data analyses. See section on consistency check.

Figure 3

Matched Pairs t Test

Changes in Spring Logic and Ethics Groups

And Pre, Post, and Follow-Up Tests



THE DIT RELATED TO BEHAVIOR IN AN EXPERIMENTAL SETTING: PROMISE KEEPING IN THE PRISONER'S DILEMMA GAME

M. KATHRYNE JACOBS

Medical College of Ohio at Toledo

Some psychologist are skeptical about the strength of the relationship between moral judgment and behavior. Mischel & Mischel (1976) for instance, emphasize the distinction between fine words and good deeds. They remind us that, "History is replete with atrocities that were justified by invoking the highest principles..." (page 107), and that the central question in moral development and behavior research remains: Do moral judgment scores merely reflect subjects' ability to invent lofty rationalizations or does moral judgment represent the way subjects generally perceive social moral situations and define the appropriate action?

There have been a number of studies showing a modest positive association between moral judgment and behavior, (Grim, Kohlberg & White, 1968; Haan, 1975; Milgram, 1963; Saltzstein, Diamond & Belenky, 1972; Schwartz, Feldman, Brown & Baingartner, 1969). While demonstrating a correspondence between judgment and behavior, the research has not clarified the nature of that relationship, particularly how behavior relates to cognitive structure at different levels of moral reasoning. (Kohlberg, 1971)

Turiel & Rothman (1972, and Rothman, 1976) investigated the behavior of young subjects presented with conflicting arguments at both one stage above and below their own stage of reasoning. Reasoning +1 was shown to influence the behavior of subjects at stage 4. This finding is an important elaboration of the research demonstrating preference for hypothetical reasoning a stage above modal stage (Rest, 1973; Rest, Turiel & Kohlberg, 1969). However in the condition where subjects were presented with a behavioral choice without supportive reasoning statements, no difference in behavior was found between

stage 3 and 4 subjects (Rothman, 1976). Rothman (1976) suggests that in the absence of a verbally presented stage-related rationale it was the situation itself that influenced subjects' behavior.

The present research investigates the effect of reasoning on behavior by examining the interaction of level of moral reasoning and the nature of the situation. The situation was specifically designed to relate to structural differences between conventional and principled levels of moral reasoning and was presented identically to all subjects. It was hypothesized that subjects would construe the situation in a manner consistent with their own cognitive developmental level and that predictable differences would emerge in behavior, decision-making and affect.

A contractual variation of Prisoner's Dilemma provided the experimental situation. Prisoner's Dilemma is a simulation game widely used in the study of conflict resolution. Subjects play a given number of trials in an effort to accumulate points having a monetary value. They may either compete or cooperate. Cooperation insures modest mutual gains while competition provides an opportunity for doubled profits and also the risk of double loss. The game therefore presents a strategic interpersonal conflict situation. In the contractual variation of Prisoner's Dilemma that conflict is resolved by a promise between subjects to cooperate (Rapaport, 1965).

In order to illicit moral reasoning, and the behavior that relates to it, it is necessary to create intrapersonal conflict, a moral dilemma. That was accomplished in the present study by introducing a promise-breaking

partner, forcing the subject to choose between keeping her promise, (cooperating and sustaining double losses), and breaking her promise, (competing to minimize losses); a situation that has no obvious solution in terms of dominant cultural standards (Kohlberg 1971). Conventional subjects were expected to break the promise while principled subjects were expected to keep it for the following reasons:

Conventional Level: While stage 3 and 4 reasoning recognizes the importance of keeping one's word, such behavior is based on a commitment to role obligation as defined by society, e.g. being a cooperative game partner or a compliant experimental subject. Faced with a promise-breaking partner, stage 3 subjects were expected to try and minimize their losses and to rationalize this behavior as being 'natural', particularly in a situation where the partner was being 'not nice'. Stage 4 subjects were expected to break their promise for similar reasons and in addition perhaps to punish the promise-breaker.

Principled Level: Stage 5 subjects were expected to keep the promise regardless of losses, because for them free agreement and contract is the binding element of obligation outside of the legal realm. The Golden Rule was expected to guide the behavior of stage 6 subjects. In addition, the agreement to behave cooperatively reflects the stage 6 valuing of human beings as ends and not means.

Assessing moral reasoning: The Defining Issues Test (DIT) purports to be a measure of general moral problem-solving strategy. It does not collect post-hoc justifications as Kolberg's measure does, but rather forces subjects to analyze, evaluate and select the issues of a moral dilemma they regard as most salient for behavior. While the DIT has been criticized as a recognition task, yielding an overestimation of reasoning capacity and thereby weakly related to behavior, Rest (1974b) proposes that the ability to recognize the central issues of a dilemma is a crucial aspect of decision making and thereby strongly related to behavior. It was partly to test this hypothesis that the DIT was chosen to assess level of moral reasoning in this study.

METHOD

Subjects. 127 adult women, ranging in age from 20 to 55, who volunteered for a study of 'decision-making patterns' were administered the DIT. 60 women who met the selection criteria (Rest, 1974a) were offered $2.00 for participating in the experiment. Participants were classified as conventional if their P%<50 and if they chose stage 3 and 4 reasoning predominantly. Assignment to the principled group was dependent on a P%≥50 and predominance of stage 5 and 6 reasoning. Only females were used in order to avoid confounding by sex, either of experimenter or confederate.

Confederate. Two women, age 25 and 31 respectively, were randomly assigned to act as partners for the subjects.

Design. A 2 x 3 randomized block design was used. Principled and conventional women were randomly assigned to one of three conditions created to produce varying levels of moral dilemma. The Cooperative condition represented a control in that subjects were not expected to experience a dilemma unless their partners broke the promise. The partners in this condition keep the promise and remained cooperative throughout the experimental procedure. In the Partial Defect condition the subject's partner broke her promise 50% of the time. In the Defect condition, the partner broke her promise on every trial. The dependent variable was the number of trials on which the subject cooperated after promising to do so. In addition, the subjects' self-reported anxiety, commitment, conflict and source of conflict were alalyzed.

Procedure. After taking the DIT and accepting an invitation to participate in the experimental aspect of the study, subjects met individually with the experimenter and another woman whom they believed to be their partner. The subject and her partner, separated by a screen which obscured their view of one another, sat side by side at a table opposite the experimenter. The rules of Prisoner's Dilemma were explained (Rapaport, 1965) and the women were instructed to signify their 'moves' by raising either their right or left hand on each trial. The right hand signified cooperation, the left competition. After each trial the women recorded their payoff points which were worth a penny apiece.

After 20 trials, during which the partner cooperated 50% of the time according to a pre-set randomized by equal pattern, a break was called. The experimenter left the room in order to allow the women to discuss the game and make strategy decisions. The partner then proposed a strategy of mutual cooperation, i.e., raising the right hand on each trial. When an unambiguous agreement was reached the experimenter returned and the game resumed, with no further communication between the women. The partner was unaware of either the moral judgment level or the condition to which the subject had been assigned. In addition to the 60 Ss whose data are reported, three principled and two conventional women refused to promise.

There were four sets of 20 trials or a total of 80 experimental trials. After 6 trials, the experimenter reported payoffs as if the partner was either continuing to cooperate, breaking the promise half the time or defecting from the promise completely. The confederate was not actually playing. Subjects were kept ignorant of how many trials were in the game to avoid the opportunity for end-game strategies.

After the first, second and third sets of trials, women responded to the brief State Anxiety form of the State-Trait Anxiety Inventory (STAI) designed by Spielburger (1970). Following the game they completed the full Trait Anxiety inventory. The Trait inventory was administered after the experiment in order to avoid an anxiety preset (Spielburger, 1970).

Subjects then responded to the following questions on a five point scale ranging from "none" to "a great deal".

1. When you and your partner agreed to cooperate, how committed did you feel to that agreement?
2. During the experiment, how much conflict did you feel about keeping the agreement?
3. How much of that conflict was due to your partner's behavior?
4. How much of that conflict was due to a desire to earn the largest payoff?
5. During the last five trials, how committed did you feel to your original cooperative agreement?

The question of obvious interest, "How much conflict was due to a desire to keep your word?" was omitted for fear of introducing experimenter expectations or producing guilt in promise-breaking women.

After formal data gathering, the full nature of the experiment was explained to each woman in an interview lasting 10-30 minutes depending on her needs and level of interest. She was encouraged to describe her thoughts and feelings during the experiment and to offer any suggestions. The interview ended at her suggestion.

RESULTS AND DISCUSSION

Promise Keeping Behavior. The 20 trials prior to the promise provided a base line of data on which to compare the groups of women. Both conventional and principled women were equally cooperative during these trials. Their mean number of cooperative responses presented in Table 1, was 10, or 50% of the trials, remarkably similar to the pre-arranged game played by the partner. It appears that in a game situation, principled women are no more cooperative by nature then conventional women. Following the promise, however, the nature of the situation changed to become one with moral implications. In the Cooperative condition with the promise-keeping partner, both principled and conventional women consistently kept their promise. The mean number of cooperative responses for both groups in this condition was 80 or 100%. No conventional or principled woman spontaneously broke the promise if their partner did not. It is evident, therefore, that the promise was sufficiently clear, meaningful and binding to have a significant effect on the cooperative behavior of both conventional and principled women.

Of greatest concern is the behavior of women in the experimental conditions; those women who agreed with the partner to cooperate and whose partner subsequently broke that promise. Their behavior was analyzed with a repeated measures analysis of variance which yielded four significant main affects. Most gratifying, there was a significant affect for level of moral judgment. Principled women, regardless of their partner's behavior, kept their promise more often than conventional women, $F(1,50) = 31.71$, $p < .001$. The major theoretical hypothesis of this study was therefore supported: Principled women's behavior was bound by a freely given agreement even in the face of a contract-breaking partner and monetary loss. The common sense hypothesis that all women would be less

promise-keeping if their partners broke the promise, was supported by the main effect for partner's behavior, $F(2,50)=137.87$, $p<.001$. Post-hoc analysis of the moral judgment X partner's behavior interaction, $F(2,50)=19.2$, $p<.001$ indicates that while the partner's promise-breaking affected all women's behavior, it tended to influence conventional women more. Conventional women were more likely to match their behavior to that of their partners. The interaction of moral judgment level with trials, $F(3,5)=2.82$, $p<.05$, suggests that over the course of the experiment, principled and conventional women displayed different patterns of cooperative responses, a fact which will be discussed in the following paragraphs.

Anxiety. The means and standard deviations of anxiety scores are presented in Table 2. Data were treated with a repeated measures analysis of variance. There were no differences in Trait anxiety between principled and conventional women. This lack of difference between the groups provides a control for the possible contention that principled subjects are more thin-skinned, compliant, neurotically moral people. The behavioral differences between principled and conventional subjects therefore cannot be explained alternatively, i.e. that principled subjects are more cooperative because they are more anxious people who dare not offend anyone.

State anxiety was investigated primarily to see if the experimentally produced dilemma had a real effect on the subject's feelings. It did. Women in the Cooperative condition experienced low levels of anxiety while women whose partners broke the promise were significantly more anxious, $F(2,50)=23.57$, $p<.001$. This effect was equally true of both conventional and principled women indicating that all women whose partners broke the promise experienced the situation as a real dilemma.

Similar to the results found for cooperative behavior, there was a significant interaction of moral judgment level and State anxiety administration at the .05 level. While this effect was not very strong and the triple interaction in both cases was non-significant an examination of the means for both promise-keeping behavior and anxiety leads to some interesting speculations. It appears that while conventional women were highly anxious during the first set of trials, their anxiety subsequently decreased as their promise-keeping decreased. Contrariwise, principled women tended to become more anxious towards the middle of the experiment when their promise-keeping was at it's lowest level. In later trials principled women increased their cooperative, promise-keeping behavior at which time their anxiety decreased. Nine of the twenty principled women with defecting partners demonstrate this particular response pattern dramatically. These behavior/anxiety patterns may be interpreted as decision making patterns involving cognitive dissonance. Both women, upon discovering their partners had broken the promise, experienced a dilemma. Both experienced heightened anxiety and several trials of indecision. Conventional women could provide themselves with no good reason for keeping the promise once their partner had broken it. They therefore adopted the most reasonable and thrifty strategy, competition. Having resolved the dilemma their anxiety dropped somewhat but not dramatically since they were concerned with the loss of money. Principled women on the other hand who initially broke their promise in response to their partner's behavior found themselves behaving incongruently with their principles. The cognitive dissonance this produced would account for the heightened anxiety. Once they returned to promise-keeping and consistent cooperative behavior, behavior and principles were again congruent and anxiety alleviated. While the data is merely suggestive at this point the idea could be further investigated by extending the number of Prisoner's Dilemma trials and increasing the number of State anxiety administrations. These speculations are further supported by the following data.

Commitment and Conflict. Responses to the questionnaire indicated that both principled and conventional women felt equally committed when they promised to cooperate. Likewise both groups of women whose partners defected felt less committed to the promise by the end of the experiment, $F(2,50)=13.6$, $p<.001$. This was particularly true in the Defect condition. However a significant effect for moral reasoning, $F(1,50)=12.79$, $p<.005$ indicated that regardless of their partner's behavior and consistent with their own behavior, principled women remained more committed to their promise than did conventional women.

Women in the Cooperative condition felt little conflict about keeping their promise while both principled and conventional women with promise-breaking partners did, $F(2,50)=11.57$, $p<.001$. The source of the conflict differed however. Conventional women indicated they were most concerned with their desire for a large payoff, $F(1,50)=11.65$, $p<.005$. The two groups were equally concerned about their partner's behavior. While the question of conflict regarding promise-keeping was not formally asked, the informal debriefing provided unexpected and significant data. Principled women spontaneously cited the contract as the rationale for their behavior more often than did conventional women, $X^2=9.27$, $p<.01$. Reflecting their stated interest in a monetary payoff, significantly more conventional women accepted the $2.00 offered to each participant in the study, $X^2=9.29$, $p<.01$.

The significant differences in anxiety, commitment, conflict and behavior indicate Ss experienced a real moral dilemma. This version of Prisoner's Dilemma is therefore a promising context within which behavior and levels of moral reasoning may be explored.

Table 1

Means and Standard Deviations of Anxiety Scores

Group	Trait Anxiety	State Anxiety Post-Contract			
		Trial 20	Trial 40	Trial 60	Pooled
Cooperative					
Principled	$\bar{X}$ 36.50	6.50	6.70	6.10	19.30
	SD 3.719	3.567	3.057	1.853	7.790
Conventional	$\bar{X}$ 37.00	7.00	6.50	6.50	20.00
	SD 7.630	2.582	2.593	2.321	7.102
Partial Defect					
Principled	$\bar{X}$ 38.20	8.00	8.00	7.70	23.70
	SD 10.737	2.708	2.789	2.452	7.181
Conventional	$\bar{X}$ 35.60	10.50	9.60	8.80	28.90
	SD 6.24	3.567	3.836	3.393	9.905
Defect					
Principled	$\bar{X}$ 36.40	9.70	10.90	9.80	30.40
	SD 5.929	3.917	5.021	4.517	13.032
Conventional	$\bar{X}$ 35.70	11.60	10.10	10.10	31.80
	SD 7.804	2.875	3.604	4.606	10.064

Table 2

Means and Standard Deviations of Anxiety Scores

Group		Trait Anxiety	State Anxiety Post-Contract			
			Trial 20	Trial 40	Trial 60	Pooled
Cooperative						
Principled	$\bar{X}$	36.50	6.50	6.70	6.10	19.30
	SD	3.719	3.567	3.057	1.853	7.790
Conventional	$\bar{X}$	37.00	7.00	6.50	6.50	20.00
	SD	7.630	2.582	2.593	2.321	7.102
Partial Defect						
Principled	$\bar{X}$	38.20	8.00	8.00	7.70	23.70
	SD	10.737	2.708	2.789	2.452	7.181
Conventional	$\bar{X}$	35.60	10.50	9.60	8.80	28.90
	SD	6.24	3.567	3.836	3.393	9.905
Defect						
Principled	$\bar{X}$	36.40	9.70	10.90	9.80	30.40
	SD	5.929	3.917	5.021	4.517	13.032
Conventional	$\bar{X}$	35.70	11.60	10.10	10.10	31.80
	SD	7.804	2.875	3.604	4.606	10.064

Social Cognition Related to Behavior in a Naturalistic Setting: A Comparison of Delinquents, Predelinquents and Nondelinquents

Edgar B. McColgan

University of Texas at Austin

(Paper presented at SRCD Convention, 1977)

The justification for studying the relationship of delinquency, or antisocial behavior, to social cognition has both empirical and pragmatic appeal. On the one hand, previous studies have indicated that adolescents who have engaged in some overt antisocial behavior are lower than nondelinquents on social cognitive development like moral judgment (e.g., Kohlberg & Freundlich, 1973; Fodor, 1972; Hickey, 1972) and social role-taking (e.g., Selman, 1973; Chandler, 1973). On the other hand, a behavioral correlate such as delinquency bridges the gap between socio-moral reasoning and moral action--a leap theorists and researchers in this area are repeatedly pressed, by pragmatists, to make. To be able to show that delinquents are deficient in their moral judgments and role-taking skills as compared with normals lends credence both to the cognitive-developmental theory on which the studies are based, and to the instruments which purport to measure special aspects of the theory.

While the "pluses" are impressive, the study of antisocial behavior is not without some drawbacks, viz., definitional problems of what counts as "antisocial," and how to come up with a control group carefully matched on variables other than social-cognitive development. I will report today on two phases of a project aimed at taking these drawbacks seriously. The general strategy of this project was to try to establish stricter comparability between antisocial and normal groups than has ever been accomplished, and to administer a battery of social-cognitive measures in order (1) to test their discriminating power for matched groups chosen on the basis of different socio-moral conduct, and (2) to study the interrelationships of the measures.

Phase I

The first phase of this study was largely a replication of previous studies dealing with social cognition and delinquency. The procedure was similar to most other studies: namely, gather data on a group of institutionalized delinquents and make some statistical comparisons with a group of "non-delinquent" subjects. Table 1 of your handout gives you some of the characteristics of the delinquent sample. Look at Column 1 of this table for the delinquent sample characteristics.

The Kohlberg MMS for this group of delinquents was about 233, which is roughly consistent with the results in previous studies with similar-aged delinquents (see McColgan & Gott, 1977, for an extensive review of the moral judgment and delinquent studies). Their understanding and interpretation of the Golden Rule was about what was expected, viz., most were unable to interpret it correctly in a specific instance. This is also generally consistent with previous studies.

Table 2 gives you some interesting data on the results using the DIT and the Moral Comprehension Measure. Look especially at the DIT score and Moral Comprehension at the bottom of the table. The Comparison subjects of Phase I were matched on age (within one year), IQ (lower portion of the average range), sex, race (all Caucasian), scoring system, and residential locale. You can see that the DIT P score is also quite sensitive to differences in the actual socio-moral behavior of these groups.

Phase II

The predelinquents of Phase II were a group of male Caucasians identified as the most disruptive and antisocial youngsters in a Contingency Management classroom in three junior high schools. The program is essentially an attempt by Community Corrections and the local Department of Special Education to provide an academic program to better meet the students' needs, and to keep them

in the mainstream of school and social life as much as possible. By having predelinquents, the effect of incarceration and institutionalization on test-taking behavior was automatically controlled. The description of these subjects suggests they display many of the same behavioral characteristics as do delinquents, with the only known exception being that they have not committed an offense for which incarceration resulted. The subjects were variously described by their teachers in the following ways: aggressive, inadequate impulse control, acts without forethought, poor self-image, excessive variations in mood, poor interactions with peers, poor work habits. In different states and at different times, many of these subjects would have gone through juvenile court proceedings rather than be maintained in school.

A nondelinquent comparison group for Phase II consisted of a matched sample from the same junior high schools. Table 1 lists some of the characteristics of these, and the predelinquent, subjects. An attempt was made to control as many variables as possible which would be likely sources of internal invalidity. Consequently, each predelinquent was matched with one control subject on 14 different dimensions, which included age, IQ, SES, race, sex, test instruments (same measures and order of presentation), interviewer, environmental conditions for all interviews, time of testing, scoring system, residential locale (same city and neighborhoods), and school. The two groups were further found to contain exactly the same number of boys from one-parent homes, and to have achieved the same average grades in school (when grades were known).

Table 3 gives you the matched-pair data for the DIT, Moral Comprehension, Kohlberg MMS, Golden Rule Task, and Chandler Role-Taking Instrument. You will notice that of the social-cognitive instruments used, none demonstrated significant differences between the groups, with one notable exception--the DIT. The antisocial group was also less able to think clearly about higher-stage concepts (Moral Comprehension) than the controls.

Now, if you will turn to Table 4, you will see that when the delinquent,

predelinquent and control groups were analyzed by ANACOVA (controlling for age, SES, and IQ) it was found that the three groups differed significantly on the DIT P score and the Kohlberg MMS. The adjusted group means on the DIT were in the predicted direction: delinquents lowest and control subjects highest (16 1/2, delinquent; 18 1/2, predelinquent; 25, control). However, when the adjusted group means were examined for the Kohlberg instrument, the delinquent subjects were found to be the most morally advanced group (227, delinquent; 194, predelinquent; 199, control)!

An examination of the intercorrelations among the social-cognitive measures across the three samples leads, at best, to confusing conclusions. There were not any strong or consistent correlations among any of the measures, although significant relationships were found between some of the variables in each group. The most striking finding was really not the correlational pattern among the social-cognitive measures: it was the Kohlberg MMS with IQ (.63, delinquent sample) and achievement (all in .60's with WRAT).

Even examining the intercorrelations in each sample led to confusing conclusions. Within an individual sample, correlations can't be explained by either common method or common construct. By common construct, we might have expected the DIT, MMS, and perhaps the Moral Comprehension Task to be similar as measures dealing with moral judgment conceived in terms of six stages, whereas the Golden Rule and Chandler Role-taking tasks are somewhat similar. The correlations did not group this way. By common method, we might have expected the Kohlberg MMS and Golden Rule, as free response measures, to be similar, as should the DIT and Moral Comprehension Measure as recognition or objective tasks. Again the correlations didn't group this way. All the correlations, when significant, were no higher than the .30's, except for the MMS and IQ and achievement, which ran in the .60's.

Conclusions

When picking up differences between institutionalized delinquents and normal adolescents matched in very imprecise ways, all the instruments confirm predictions. But, when the two groups are brought closer together (i.e., the delinquents are not the worst offenders and the controls are carefully matched) the discriminative power of all the instruments, except the DIT, is no longer present. How can we explain the divergent results of Phases I and II? Let me briefly offer some tentative possibilities.

(1) For one thing, it is clear that a person's ability to appreciate Principled thinking is a more sensitive discriminator of fine differences between groups differing in socio-moral conduct than is any other instrument used, including the Kohlberg instrument.

(2) The DIT, as a recognition/comprehension task, is not as heavily loaded on verbal expressive skills as are the other instruments, particularly the Kohlberg instrument. This fact was supported by correlations with IQ and school achievement. I believe the data support the suggestion that a person's tacit awareness of principled arguments has a lot to do with one's decision regarding overt action. I'm less convinced that being able to verbally explain arguments has much to do with people's actions.

(3) The Kohlberg MMS as a way of indexing moral development may not be the best way of indexing a person's development. As Rest has suggested (Rest, 1977) we might do well to look at the total range of a subject's distribution of responses across the six stages rather than attempt to categorize each subject by his/her modal stage reasoning.

(4) It could also be that the different instruments are either not measuring the same construct, or are measuring the construct "social cognition" in different ways. This would be consistent with Carolyn Shantz's (Shantz, 1975) suggestion that social cognition is not a unitary construct.

(5) Perhaps there is really no difference between the predelinquent and

control groups in this study on how they judge socio-moral conflicts. Rather, the differences found with the DIT may be due to covariation with some yet uncontrolled variable. The burden of proof for this possibility rests on replication by other well-designed studies.

At the very least, Phase II demonstrated that social-cognitive instruments are not as potent discriminators as the literature would suggest. This study raises some serious questions about the reliability of present assessment methods with delinquents, and indicates that more research needs to be done with instruments that can discriminate characteristics of antisocial adolescents--instruments like the DIT.

Table 1

Subject Characteristics: Phases I and II

Characteristic	Phase I	Phase II	
	Delinquent	Predelinquent	Control
Age (months)			
M	193.97	167.81	170.35
SD	11.89	12.34	8.26
IQ			
M	93.86	94.88	95.08
SD	8.73	15.18	14.26
SES			
M	26.69	37.77	37.96
SD	22.50	27.52	27.29

Note: Delinquent sample $n = 29$; Predelinquent and Control samples $n = 26$.

Table 2

Means, Standard Deviations, and t Tests on DIT Stage Scores and Moral Comprehension: Phase I

Variables	Groups					
	Delinquent		Comparison			
	M	SD	M	SD	t test	p value
DIT Stage Scores						
2	10.57	3.37	9.72	3.53	-1.000	.321
3	26.15	4.71	22.88	5.96	-2.440	.017
4	36.90	6.13	30.65	7.01	-3.830	<.001
5A	12.76	3.66	18.25	6.04	4.316	<.001
5B	3.74	2.03	6.38	2.92	4.157	<.001
6	2.30	1.80	4.07	2.24	3.491	<.001
A	2.93	2.42	4.07	2.79	1.759	.083
M	4.66	2.11	3.98	1.76	-1.453	.151
P Score	18.79	5.16	28.70	7.39	6.163	<.001
Moral Comprehension	16.03	3.60	12.70	3.56	-3.801	<.001

Note: Delinquent sample $n = 29$; Comparison sample $n = 41$.

Table 3

Means, Standard Deviations, and Matched Pair *t* Tests on Dependent Measures: Phase II

Variables	Groups					
	Predelinquent		Control			
	M	*SD*	*M*	*SD*	*t* test	*p* value
DIT Stage Scores						
2	11.67	3.18	10.51	3.16	-0.66	.513
3	29.55	5.89	24.62	6.28	-1.67	.107
4	32.56	5.29	33.46	5.35	0.36	.720
5A	12.76	4.68	16.41	5.63	2.27	.032
5B	2.05	1.53	3.85	2.92	1.90	.070
6	2.31	2.37	3.46	2.00	1.27	.216
A	5.32	3.09	3.40	2.01	-1.41	.172
M	3.97	2.70	4.17	2.18	0.17	.866
P Score	16.86	4.64	23.72	7.02	3.58	<.001
Moral Comprehension	15.00	2.68	16.69	2.49	2.64	.014
Kohlberg Moral Maturity Score	189.27	13.97	196.08	30.58	1.13	.269
Piaget Golden Rule Task	223.08	32.34	228.85	37.88	0.72	.478
Chandler Role Taking Task	12.77	4.11	12.12	5.05	-0.54	.592

Note: $n = 26$ in both groups.

Table 4

F Ratios and *p* Values for Dependent Measures: Samples from Phases I and II Combined

Variables	ANOVA		ANACOVA	
	F	p value	*F*	p value
DIT Stage Scores				
2	.36	.697	.75	.473
3	1.89	.158	2.34	.103
4	1.66	.197	1.08	.344
5A	1.93	.152	2.35	.040
5B	1.94	.151	1.44	.244
6	1.01	.371	.88	.417
A	2.41	.097	2.65	.077
M	.23	.795	.17	.847
P Score	3.65	.031	4.79	.011
Moral Comprehension	2.12	.127	2.04	.138
Kohlberg Moral Maturity Score	21.67	<.001	6.27	.003
Piaget Golden Rule Task	.39	.677	.80	.451

Note: $n = 81$.

Indexing Moral Development

Mark L. Davison and Stephen Robbins

University of Minnesota

Kohlberg (1969, 1971) has outlined a six stage theory of moral development. Each successive stage in the theory is characterized by moral reasoning which is more complex, more comprehensive, more integrated, and more differentiated than the reasoning of earlier stages. According to Kohlberg's theory the child develops by movement through a sequence of steps with the following bases of moral order: internal compulsion and power (Stage 1), interpersonal exchange and need satisfaction (Stage 2), social expectations (Stage 3), authroity figures (Stage 4), a legitimate social contract (Stage 5), and individual conscience (Stage 6).

Kohlberg himself has developed an interview procedure for assessing an individual's level of moral development. Rest (Note 1, Rest, Cooper, Coder, Masanz, and Anderson, 1974) has developed an objective measure of moral level. The test consists of six stories describing moral dilemmas. After reading each story, the subject is first asked to answer a yes-no question indicating how s/he thinks the central character of the story ought to respond to the dilemma. Second, the respondent is asked to rate each of several issues on a five-point scale of importance in deciding what ought to be done. And finally, the respondent is asked to rank order the four issues which s/he thinks are the most important. Each of the issues reflects reasoning characteristic of either Stage 2, 3, 4, 5A, 5B, or 6 in Kohlberg's theory.

Because Rest's issues can be divided into six sets, one for each of the stages; 2, 3, 4, 5A, 5B and 6; in Kohlberg's theory, the test readily yields

x_{ij} = the rating given by subject $\underline{i}$ to issue $\underline{j}$.

$\bar{x}_{i\cdot} = \frac{1}{n} \Sigma_j x_{ij}$, where $\underline{n}$ is the number of issues.

$\bar{x}_{j\cdot} = \frac{1}{N} \Sigma_i x_{ij}$, the mean response to item j in a standardization sample of N = 1080.

$\bar{x}_{\cdot\cdot} = \frac{1}{Nn} \Sigma_i \Sigma_j x_{ij}$, the grand mean of responses in the standardization sample.

D itself is a weighted sum of double standardized item responses where the weights were empircally derived via a scaling algorithm described by Schönemann (1970). (Davison (1978) and Davison, Robbins, and Swanson (Note 2) supply the justification for adopting Schönemann's algorithm.) While the weights vary from item to item, the average weight given to Stage 2 items is lower than the average weight given to Stage 3 items which in turn is lower than the average weight given to Stage 4 items etc.

Because it uses information from items keyed to every stage, D is more intuitively appealing than P as a measure of overall development. The most important comparison between the indices, however, is the comparison of their empirical properties. We will now summarize these comparisons. With one exception to be noted below, none of the following data come from the sample on which D was standardized.

Methods and Results

Reliability

In a sample of 160 students, 40 junior high, 40 senior high, 40 college, and 40 graduate students, the internal consistency reliability estimate for

P was .77. The comparable estimate for D was .79. The intercorrelation between the two indices was .74.

In a sample composed of 33 college students and 19 ninth graders retested after a two to three week interval, the test-retest correlations for P and D respectively were .77 and .76 respectively. Within the ninth graders, the test-retest correlations were .81 and .92. For the college students, the figures were .70 and .68. None of these differences between test-retest reliabilities for P and D are statistically significant at the .05 level. In a final sample composed of students and adults retested after an interval of eight to eighteen weeks, the test-retest reliabilities were .82 and .87. Again, the difference is not significant at the .05 level.

Validity

In three different samples, the correlations between the two indices were in the .70's. As cited above, in the heterogeneous sample composed of junior high, senior high, college, and graduate students, the correlation between the two indices was .74. A correlation of .72 was obtained in a sample of 20 composed primarily of education and psychology undergraduates. In the standardization sample of 1080 subjects, the correlation between P and D was .78.

In the sample of 160 junior high, senior high, college, and graduate students, males and females did not differ significantly in either their P or D scores. Nor were P and D significantly related to SES as measured by father's education. Both P and D were modestly correlated with SES as measured by father's occupation ($r_P = .20$, $p < .05$, $r_D = .28$, $p < .05$).

Table 1 shows that both P and D display a roughly similar pattern of correlations with measures of cognitive ability (Differential Aptitude Test Composite Score), comprehension of moral issues, political tolerance, and

law and order orientation. Basically, both indices are more highly correlated with measures associated with moral development than with the measure of general cognitive ability. None of the differences in Table 3 are statistically signifcant.

Longitudinal Analyses

The major difference between the indices seems to be that D is more sensitive to naturally occurring longitudinal change. As Table 2 shows, both P and D changed significantly from 1972-1976 in a sample of 54 junior high and high school students, but as reflected by the associated F statistics, the trend for D was stronger than that for P.

Similar differences occurred in the 21 subjects tested by Elaine Wilson in 1974 and 1976. In this sample, however, only D changed significantly. Examination of the subject's stage scores revealed that only Stage 3 scores changed significantly, a stage reflected in D scores but not by P scores.

Discussion

Though there were fluctuations from sample to sample, the reliabilities for the two indices were generally in the .70's or .80's. The correlations between the two indices tended to be in the low .70's. Neither index varied significantly as a function of sex or father's education, though D was modestly related to father's occupation in one sample. Both indices displayed a similar pattern of correlations with outside measures. The major difference between P and D seemed to be that D was more sensitive to naturally occurring longitudinal change.

D is a complex index to compute by hand, because it is a weighted sum of transformed item rating scores. One wonders whether all of the

complexity is necessary. Would we do just as well by taking an unweighted sum of item socres? If a weighted sum is necessary, need we use empirically derived item weights as with D or could we give each item a weight equal to its corresponding stage? While we have not yet examined the simpler alternatives as thoroughly as P and D, a simple item sum (after reverse scoring responses to nonprincipled issues) has proven somewhat less reliable, less highly correlated with outside measures, and less sensitive to longitudinal trends than D. More work with the simple item sum is needed, however. As yet we have no data to report on weighting each item by its corresponding stage number.

Because D incorporates information from responses to all stages, it has more logical appeal than P as an overall measure of development. More importantly, however, incorporating information from responses to all issues has yielded an index seeningly more sensitive to longitudinal change. When significant longitudinal stage change was observed only for Stage 3, a stage not reflected by P, D (but not P) changed significantly. In a second longitudinal sample where significant stage change occurred in both Principled and Nonprincipled Stages, both measures changed significantly, but the trends in D were stronger than those in P. For those interested in assessing overall change in a manner which reflects lower stage as well as upper stage change, D would seem to be the superior index of development. Because both P and D are somewhat complex to compute, we now offer a computerized scoring service which provides stage scores, P scores, and D scores for each subject as computed from their DIT responses. Those who wish to use the service should contact either Drs. James R. Rest or Mark L. Davison, 330 Burton Hall, Department of Social, Psychological, and Philosophical Foundations of Education, University of Minnesota, Minneapolis, Minnesota 55455.

TABLE 1

Correlations of D and P with measures of cognitive ability (DATVN), comprehension of moral issues (COMP), law and order orientation (LO), and political tolerance (PT).

	DATVN	COMP	LO	PT
P	.42*	.63*	-.52*	.58*
D	.47*	.63*	-.49*	.55*

TABLE 2

Longitudinal Trends in P and D

Rest Data

	1972		1974		1976		F
	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	
P	20.32	7.39	21.74	7.86	26.10	7.72	20.06**
D	-.06	.37	.24	.48	.32	.46	24.86**

Wilson Data

	1974		1976		t
	$\bar{x}$	s	$\bar{x}$	s	
P	20.33	10.55	22.15	7.21	-.97
D	-.04	.56	.21	.50	-2.64*

*P < .05

**P < .01

Addendum:

Personality Correlates of the DIT

Since stages of moral judgment are just one aspect of a person's personality, many interesting questions arise about the relation of moral judgment to the rest of personality: How closely tied or independent is moral judgment to other aspects of a person's social development such as autonomy, self esteem, role taking ability, ego identity, etc.? Are certain personality characteristics prerequisites to high moral judgment development--such as liberalism, intellectualism, social sensitivity? Is moral judgment development a prerequisite to general personality development and/or integration, as theorists such as Loevinger and Erikson suggest? Could we improve the predictability of behavior by combining information about moral judgment along with information about other personality characteristics, such as ego strength, empathy, decisiveness, independence, etc.? Does the person who uses Stage 2 predominantly tend to be more machiavellian? Does the person who uses Stage 3 tend to be more socially conforming and approval seeking? Does the person who uses Stage 4 tend to be more authoritarian, generally?

Questions such as these are addressed by the studies summarized below. It is difficult to summarize the results since so many different variables have been explored and since many relationships are unclear as yet. Nevertheless this information might be useful to researchers planning future research in the area, and I wanted to round out this summary of DIT research by citing, however briefly, this work.

Personality Correlates of DIT

Study and sample	Personality Test and Variable	Correlation
Allen & Kickbusch, 1976	Rosenberg scale of self-estemm	.02
430 9th graders	Rosenberg scale of faith in people	.02
	Jourard: self disclosure	
	- on attitude and value	-.06
	- on personality	-.02
	Sense of power in school	.01
	Generalized attitudes toward school	.04
Blackner, 1975	Tennessee Self Concept Scale	
80 high school Ss	- positive total scale (self esteem)	.15
	- moral ethical self	.19
	- personality integration	-.03
Blackner, 1975	Tennessee Self Concept Scale	
80 adults	- positive total scale (self esteem)	-.20
	- moral ethical self	-.12
	- personality integration	.05
Bloomberg, 1974	Rotter's I-E scale	-.06
53 undergrads		
Cauble, 1976	Constantinople's Inventory of Pers. Dev.	.01
90 college Ss	Marcia's Ego Identity Statuses	NS
Dortzbach, 1975	Rotter's I-E scale	-.17
185 adults		
Hartwich, 1975	California Personality Inventory	
98 undergraduates	- Dominance	.11
	- Capacity for status	.32
	- Sociability	.11
	- Social presence	.17
	- Self acceptance	.20
	- Sense of well-being	.20
	- Responsibility	.33
	- Socialization	-.05
	- Self control	.18
	- Tolerance	.39
	- Good impression	.09
	- Communality	.15
	- Achievement via conformance	.25
	- Achievement via independence	.48
	- Intellectual efficiency	.42
	- Psychological mindedness	.32
	- Flexibility	.11
	- Femininity	.03
Johnson, 1974	Intrinsic-extrinsic religious orientation	NS
66 high school Ss		

Study and sample	Personality Test and Variable	Correlation
Masanz, 1975	Teachers; ratings	
33 high school Ss	- Initiative, industry	.40
	- Motivation	.35
	- Cooperation	NS
	- Emotional stability	NS
	- Common sense	NS
	- Leadership	NS
	- Personality	NS
	- Reliability	NS
	- Concern for others	NS
	- Honesty	NS
	- Respect for authority	NS
	- Work effort	NS
McColgan, 1975	Jesness Test	
	- Social maldjustment	.20
	- Value orientation	.28
	- Immaturity	-.18
	- Autism: distort reality	-.03
	- Alienation: distrust of others	.36
	- Manifest aggression	.06
	- Withdrawal	.17
	- Social anxiety	.29
	- Repression	-.34
	- Denial to acknowledge conflict	-.33
	- Antisocial index	.14
	I-Level	NS
Morrison, Toews & Rest, 1973	Minnesota Affect Assessment	
73 junior highs	- General school interest	.16
	- Autonomy	-.11
	- Self expression	-.21
	- Academic set	-.05
	- Fine arts	-.03
	- School personnel	.00
	- Importance of affect	-.15
	- Intrinsic motivation	.13
	- External locus of control	.21
	- Need for direction	.09
Sanders, 1976	Norwicki-Strickland's Locus of Control	.30
42 junior highs		
Schneeweis, 1974	Allport-Vernon-Lindzey Study of Values	
64 high school Ss	- Theoretical	-.04
	- Economic	-.07
	- Aesthetic	-.03
	- Social	-.02
	- Political	.05
	- Religious	.07

Study and sample	Personality Test and Variable	Correlation
Schomberg, 1975	Omnibus Inventory	
35 college Ss	-Thinking introversion	.36
	- Complexity	.45
	- Autonomy	.47
	- Practical outlook	-.51
	- Masculinity/femininity	-.11
	- Theoretical orientation	.29
	- Estheticism	.32
	- Religious orientation	.10
	- Social extroversion	.11
	- Impulse expression	.19
	- Personal integration	.31
	- Anxiety level	.39
	- Altruism	.26
Standring, 1976	Rokeach Value Survey	
33 college Ss	- Comfortable life	.05
	- Exciting	.38
	- Sense of accomplishment	.23
	- World at peace	.30
	- World of beauty	.16
	- Equality	.36
	- Family security	-.37
	- Freedom	.50
	- Happiness	-.24
	- Inner harmony	.34
	- Mature love	-.05
	- National security	-.52
	- Pleasure	.63
	- Salvation	-.60
	- Self-respect	.18
	- Social recognition	-.13
	- True friendship	.04
	- Wisdom	.12
Troth, 1974	Gurin's Internal-external control	
47 college Ss	- Total	.12
	- Control ideology	.02
	- Personal control	.21
	- System modifiability	.15
Winocur & Rogers, 1975	Machiavellianism (Mach V)	.01
87 undergraduates	Rotter's I-E scale	-.14
	Ring & Wallston's Performance Style Typology	
	- P: wants to be own person	-.11
	- T: enjoys interpersonal games	.20
	- C: "social chameleon"	-.31

References

Allen, R. An evaluation of the Nicolet High School Confluent education project for the second year, 1974-75. Unpublished manuscript, Nicolet High School, Glendale, Wisconsin, 1976.

Alozie, C.F. An analysis of the interrelationship of two measures used in the measurement of moral judgment development: the Kohlberg Moral Judgment Interview and The Rest Defining Issues Test. Unpublished doctoral dissertation, University of Minnesota, 1976.

Alston, W.P. Comments on Kohlberg's "From is to ought." In T. Mischel (Ed.), Cognitive development and epistemology, New York: Academic Press, 1971, 269-284.

Anderson, S.M. The assessment of moral judgment development for dental hygiene education. Unpublished Master's Plan B Paper, University of Minnesota, 1975.

Armstrong, P.M. Psychological comparison of prison inmates and parolees. Unpublished doctoral dissertation, Southern Illinois University, 1975.

Balfour, M.J. An investigation of a school-community involvement program's effect on the moral development of its participants. Unpublished master's paper, University of Minnesota, 1975.

Biskin, D. Personal Communication, August 28, 1974.

Blatt, M. Studies on the effects of classroom discussion upon children's moral development. Unpublished doctoral dissertation, University of Chicago, 1970.

Blatt, M. & L. Kohlberg. The effects of classroom moral discussion upon children's level of moral judgment. Unpublished manuscript, Harvard University, 1973.

Blackner, G.L. Moral development of young adults involved in weekday religious education and self-concept relationships. Unpublished doctoral dissertation, Brigham Young University, 1975.

Bloom, R.B. Morally speaking, who are today's teachers? Phi Delta Kappan, 1976, 57, 624-625.

Bloomberg, M. On the relationship between internal-external control and morality. Psychological Reports, 1974, 35, 1077-1078.

Bracht, G.H. and Glass, G.V. The external validity of experiments. American Educational Research Journal, 1968, 5, 437-474.

Bransford, C. Moral development in college students. Unpublished manuscript, St. Olaf College, 1973.

Briskin, A.S. Relationships between three stage-level models of development. Unpublished paper, University of Minnesota, 1975.

Campbell, D.T. and Stanley, J. C. Experimental and Quasi-experimental research on teaching. In N. L. Gage, Handbook of Research on Teaching. Chicago: Rand McNally, 1963, 171-246 and Reprinted R. McN., 1966.

Carroll, J. & Rest, J. R. Study in progress, University of Minnesota, 1977.

Cauble, M.A. Formal operations, ego identity, and principled morality: are they related? Developmental Psychology, 1976, 12, 363-364.

Chandler, M.J. Egocentrism and antisocial behavior: The assessment and training of social perspective-taking skills. Developmental Psychology, 1973, 9(1), 1-6.

Coder, R. Moral judgment in adults. Unpublished doctoral dissertation, University of Minnesota, 1975.

Copa, B.A. Relationships between theories of conceptual and moral development. Unpublished master's dissertation, University of Minnesota, 1975.

Cooper, D. The analysis of an objective measure of moral development. Unpublished doctoral dissertation, University of Minnesota, 1972.

Davison, M.L. On a unidimensional, metric unfolding model for attitudinal and developmental data. Psychometrika, 1978 (in press).

Davison, M.L., Robbins, S., & Swanson, D.B. Stage structure in objective moral judgments. Unpublished manuscript, University of Minnesota, 1976 (Reference Note 2).

Davison, M.L. & Robbins, S. Indexing moral development. Unpublished manuscript, University of Minnesota, 1977.

Davison, M.L., Robbins, S., & Swanson, D. A Fortrand Program for scoring the Defining Issues Test. Unpublished manuscript, University of Minnesota, 1977.

Dispoto, R. Socio-moral reasoning and environmental acitivy, emotionality, and knowledge. Unpublished doctoral dissertation, Rutgers University, 1974.

Dortzbach, J.R. Moral judgment and perceived locus of control: a cross-sectional developmental study of adults, aged 25-74. Unpublished doctoral dissertation, University of Oregon, 1975.

Erickson, V.L., S. Colby, P. Libbey, & G. Lohman. The young adolescent: a curriculum to promote psychological growth. In G.D. Miller (Ed.), Developmental Education. St. Paul, Minnesota: Minnesota Department of Education, 1976.

Ernsberger, D.J. Intrinsic-extrinsic religious identification and level of moral development. Unpublished doctoral dissertation, University of Texas, 1976.

Fodor, E. Delinquency and susceptibility to social influences among adolescents as a function of level of moral development. Journal of Social Psychology, 1972, 86, 257-260.

Froming, W. & R.G. Cooper. Predicting compliance behavior from moral judgment scales. Unpublished manuscript, University of Texas, 1976.

Froming, W.J. & McColgan, E. A comparison of two measures of moral judgment. Unpublished manuscript, University of Texas at Austin, 1977.

Gallia, T.J. Moral reasoning in college science and humanities students: summary of a pilot study. Unpublished manuscript, Glassboro State College, 1976.

Gibbs, J.C. & C. J. Fedoruk. Experimentally induced change in moral judgment. Paper presented at the Symposium of the Jean Piaget Society, Philadelphia, 1975.

Grim, P.F., Kohlberg, L. & White, S.H. Some relationships between conscience and attentional processes. Journal of Personality and Social Psychology, 1968, 8, 239-252.

Gutkin, D.C. & J. M. Suls. Relationship between principled moral reasoning, internal-external control, and the ethics of personal conscience. Unpublished manuscript in preparation, State University of New York at Albany, 1976.

Guttenberg, R. Videotaped moral dilemmas: altering the presentation of the stimuli in the Defining Issues Test. Unpublished manuscript, Brown University 1975.

Haan, N. Hypothetical and actual moral reasoning in a situation of civil disobedience. Journal of Personality and Social Psychology, 1975, 32, 2, 255-270.

Haan, N., M.B. Smith, & J. H. Block. The moral reasoning of young adults: political-social behavior, family background, and personality correlates. Journal of Personality and Social Psychology, 1968, 10, 183-201.

Hartwick, R. Personal Communication, June 6, 1975.

Hays, W.L. Statistics for psychologists. New York: Holt, Rinehart & Winston, 1963.

Hickey, J. The effects of guided moral discussion upon offenders' level of moral judgment. Unpublished doctoral dissertation, Boston University, 1972.

Holstein, C.B. Irreversible, stepwise sequence in the development of moral judgment: a longitudinal study of males and females. Child Development, 1976, 47, 51-61.

Hudgins, W., & Prentice, N. M. Moral judgment in delinquent and nondelinquent adolescents and their mothers. Journal of Abnormal Psychology, 1973, 82(1), 145-152.

Hurt, B.L. Psychological education for college students: a cognitive-developmental curriculum. Unpublished doctoral dissertation, University of Minnesota, 1974.

Ismail, M.A. A cross-cultural study of moral judgments: the relationship between American and Saudi Arabian university students on the Defining Issues Test. Unpublished doctoral dissertation, Oklahoma State University, 1976.

Jacobs, M.K. Women's moral reasoning and behavior in a contractual form of prisoners' dilemna. Paper presented at AERA, San Francisco, California, 1976.

Johnson, M.A. A study of relationships between religious knowledge, moral judgment and personal religious orientations. Unpublished manuscript, Temple University, 1974.

Keasey, C.B. The lack of sex differences in the moral judgments of pre-adolescents. The Journal of Social Psychology, 1972, 86, 157-158.

Keller, B.B. Verbal Communication characteristics of couples at principled, conventional or mixed levels of moral development. Unpublished Master's dissertation, College of William and Mary, 1975.

Kickbusch, K. Personal communication, 1975.

Kohlberg, L. The development of modes of moral thinking and choice in the years 10 to 16. Unpublished doctoral dissertation, University of Chicago, 1958.

Kohlberg, L. Stage and sequence: the cognitive-developmental approach to socialization. In D. Goslin (Ed.) Handbook of socialization theory and research. Chicago: Rand McNally, 1969, 347-480.

Kohlberg, L. From is to ought: How to commit the naturalistic fallacy and get away with it in the study of moral development. In T. Mischel (Ed.), Cognitive Development and Epistemology. New York: Academic Press, 1971.

Kohlberg, L., & Freundlich, D. The relationship between moral judgment and delinquency. Unpublished manuscript, Harvard University, 1973.

Kramer, R. Moral development in young adulthood. Unpublished doctoral dissertation, University of Chicago, 1968.

Krause, D. Personal communication, August 22, 1974.

Kuhn, D. Short term longitudinal evidence for the sequentiality of Kohlberg's early stages of moral judgment. Developmental Psychology, 1976, 12, 162-166.

Leming, J.S. Ego and conscience factors as determinants of resistance to cheating among college undergraduates. Unpublished manuscript, State University of New York at Stony Brook, 1976.

Loevinger, J. Ego development. San Francisco, Josey-Bass, 1976.

Lockwood, A.L. The effects of values clarification and moral development curricul on school-age subjects: a critical review of recent research. Unpublished manuscript, University of Wisconsin, 1977.

Martin, R. The reliability and validity of the Defining Issues Test. Developmental Psychology, 1977, in press.

Masanz, J.A. Moral judgment development in adolescent girls. Unpublished master's dissertation, University of MInnesota, 1975.

Meyer P.G. Intellectual development of college students as measured by analysis of religious content. Unpublished doctoral dissertation, University of Minnesota, 1975.

McColgan, E. Social cognition in delinquents, pre-delinquents and non-delinquents. Unpublished doctoral dissertation, University of Minnesota, 1975.

McColgan, E.B. Social cognition in delinquent, predelinquent and nondelinquent adolescents. Unpublished manuscript, University of Texas at Austin, 1977.

McColgan, E.B. & Gott, S. Delinquency as a correlate of moral judgment: A review of the literature. Unpublished manuscript, University of Texas at Austin, 1977.

McGeorge, C. The susceptibility to faking of the Defining Issues Test of moral development. Developmental Psychology, 1975, 11, 108.

McGeorge, C. Some correlates of principled moral thinking in young adults. Unpublished manuscript, University of Canterbury, 1975.

Mischel, W. & Mischel, H. N. A cognitive social learning approach to morality and self-regulation. In T. Lickona (Ed.), Moral development and behavior, New York: Holt, Rinehart & Winston, 1976, Chapter 4.

Milgram, S. Behavioral study of obedience. Journal of Abnormal and Social Psychology, 1963, 67, 371-378.

Morrison, T., O. Toews, & J. Rest. An evaluation of a jurisprudential model for teaching social studies to junior high school students. Study in progress, University of Manitoba, 1973.

Mullen, P. Personal Communication, November 5, 1975.

Nelson, K.H. The Sierra project: from theory to research. Paper presented to the Moral Development and Psychological Education Conferences, 1976.

Nicholls, J.G. Effects of moral maturity and perception of effort on achievement evaluation. Unpublished manuscript, Victoria University of Wellington, New Zealand, 1976.

Panowitsch, H.R. Change and stability in the Defining Issues Test. Unpublished doctoral dissertation, University of Minnesota, 1975.

Piaget, J. [The moral judgment of the child] (M. Gabain, trans.). New York: The Free Press, 1965. (Originally published, 1932.)

Piwko, J. The effects of a moral development workshop. Unpublished manuscript, 1975.

Rapaport, A. and Chammah, A. M. Prisoner's Dilemma: A study in conflict and cooperation. Ann Arbor: The University of Michigan Press, 1965.

Rest, J. The hierarchical nature of moral judgment. Journal of personality, 1973, 41, 86-109.

Rest, J.R. The cognitive-developmental approach to morality: the state of the art. Counseling and Values, 1974a, 18, 64-78.

Rest, J. Manual for the Defining Issues Test: An objective test of moral judgment development. Available from author (330 Burton Hall, University of Minnesota, Minneapolis, Minnesota 55455), 1974b.

Rest, J.R. Developmental psychology as a guide to value education: a review of "Kohlbergian" programs. Review of Educational Research, 1974c, 44, (2), 241-259.

Rest, J. The validity of tests of moral judgment. In J. Meyer, B. Burnham, J. Cholvat (Eds.), Value education: theory, practice, problems, prospects. Waterloo, Ontario, Canada: Wilfrid Laurier University Press, 1975a, 103-116.

Rest, J. Longitudinal study of the Defining Issues Test: a strategy for analyzing developmental change. Developmental Psychology, 1975b, 11, 738-748.

Rest, J. New approaches in the assessment of moral judgment. In T. Lickona (Ed.), Moral development and behavior. New York: Holt, Rinehart & Winston, 1976a, 198-220.

Rest, J.R. Moral judgment related to sample characteristics. Unpublished manuscript, University of Minnesota, 1976b.

Rest, J.R. A theoretical analysis of moral judgment development. Unpublished manuscript, University of Minnesota, 1977a.

Rest, J.R. The stage concept in moral judgment research. Unpublished manuscript, University of Minnesota, 1977b.

Rest, J.R. Longitudinal study, in progress. University of Minnesota, 1977c.

Rest, J., Turiel, E., and Kohlberg, L. Relations between level of moral judgment and preference and comprehension of the moral judgment of others. Journal of personality, 1969, 37, 225-252.

Rest, J.R., Ahlgren, C., Mackey, J. Minneapolis Police Report. Unpublished manuscript, University of Minnesota, 1972.

Rest, J.R., Cooper, D., Coder, R., Masanz, J., and Anderson, D. Judging the important issues in moral dilemmas--an objective test of development. Developmental psychology, 1974, 10, (4), 491-501.

Rest, J.R. & B. Feldman. Correlations of the DIT with political attitudes. Study in progress, University of Minnesota, 1973.

Rothman, G. The influence of moral reasoning on behavioral choices. Child Development, 1976, 47, 397-406.

Saltzstein, H.D., Diamond, R. M. & Balenky, M. Moral judgment level and conformity behavior. Developmental Psychology, 1972, 7, 327-336.

Sanders, N. Pretesting in the pilot study schools of the skills for ethical action project. Unpublished manuscript, Research for Better Schools, Inc., 1976.

Selman, R.L. The relation of role-taking and moral judgment stages: A theoretical and empirical analysis. Unpublished manuscript, Harvard University, 1973.

Schaie, K., G. Labouire, & B. Buech. Generational and cohort-specific differences in adult cognitive functioning: a fourteen-year study of independent samples. Developmental Psychology, 1973, 9, 151-166.

Schneeweis, T.G. The relationship between the Allport-Vernon-Lindzey study of values and an objective measure of moral judgment. Unpublished master's dissertation, Moorhead State College, 1974.

Schomberg, S.F. Some personality correlates of moral maturity among community college students. Unpublished manuscript, University of Minnesota, 1975 (in progress).

Schomberg, S.F. The interpretation of college student's ratings on selected personality variables within the context of the developmental variable of moral judgment. Unpublished doctoral dissertation, University of Minnesota, 1977 (in progress).

Schomberg, S.F. and E. Balkcum. Evaluation of two instructional approaches in teaching ethics. Study in progress, University of Minnesota, 1976.

Schomberg, S.F. & J. Nelson. Evaluation of a Christian Ethics Course for seminarians. Study in progress, University of Minnesota, 1976.

Schönemann, P.H. On metric multidimensional unfolding. Psychometrika, 1970, 35, 349-366.

Schwartz, S.H., Feldman, K. A., Brown, M. E. & Heingartner, A. Some personality correlates of conduct in two situations of moral conflict. Journal of Personality, 1969, 37, 41-57.

Shantz, C. The development of social cognition. In E. M. Hetherington (Ed.), Review of Child Development Research, Vol. 5. Chicago: University of Chicago Press, 1975, Chapter 6.

Siegal, M. An experiment in moral education: AVER in Surrey. Paper presented at Annual Conference, Canadian Society for the Study of Education, Toronto, Ontario, 1974.

Snow, R. Representative & Quasi - Representative Designs for research on teaching. Review of Educational Research, 1974, Vol. 44, No. 3, 265-292.

Spielburger, C.D., Goruch, R. L., & Lushene, R. E. Manual for the State-Trait Anxiety Inventory. Palo Alto, California: Consulting Psychologist Press, 1970.

Sprecher, P. Moral judgment in pre-adolescents: peer morality versus authority morality. Unpublished master's dissertation, University of Wisconsin, 1976.

Sprinthall, N.A. and Bernier, J. E. Moral and cognitive development for teachers: a neglected area. Chapter for Fordham University Symposium: Programs and Rationale in Value - Moral Education, 1977.

Standring, D.E. Values, value change and attitudes to occupations as a function of moral judgment level. Unpublished master's thesis, Victoria University of Wellington, New Zealand, 1976.

Tracy, J.J. and Cross, H. J. Antecedents of shift in moral judgment. Journal of Personality and Social Psychology, 1973, 26, 238-244.

Troth, A.G. An assessment of impacts on students of the "Values" course. Unpublished manuscript, St. Olaf College, 1974.

Turiel, E. An experimental test of the sequentiality of developmental stages in the child's moral judgments. Journal of Personality and Social Psychology, 1966, Vol. 3, No. 6, 611-618.

Turiel, E. and Rothman, G. The influence of reasoning on behavioral choices at different stages of moral development. Child Development. 1972, 43, 741-756.

Walker, L. The effect of a narrative model on stages of moral development. Unpublished manuscript, University of New Brunswick, Canada, 1974.

Weber, R.G. The nature of authoritarianism and its relationship to other personality variables and stages of moral judgment. Unpublished doctoral dissertation, University of Minnesota, 1974.

Weisbroth, S.P. Moral judgment, sex, and parental identification in adults. Developmental Psychology, 1970, 2, 396-402.

White, C.B. Moral judgment in college students: the development of an objective measure and its relationship to life experience dimensions. Unpublished doctoral dissertation, University of Georgia, 1973.

White, C. Moral development in Bahamian school children: a three year examination of Kohlberg's stages of moral development. Unpublished manuscript, Southern Methodist University, 1977.

Whiteley, J.M. The scope and purpose of the UCI residential education project. Unpublished manuscript, University of California, Irvine. 1976.

Whiteley, J.M. & K. H. Nelson. Personal communication, June 27, 1976.

Winocur, S. & R. S. Rogers. Individual differences, locus of issue and the resolution of moral dilemmas. Unpublished manuscript, University of Reading, 1975.

Wohlwill, J.F. The study of Behavioral Development. New York: Academic Press, 1973.

Yussen, S.R. Moral reasoning from the perspective of others. Child Development, 1976, 47, 551-555.

END